Transposon Tagging in Strawberry and Potato and Characterization of Representative Strawberry Mutants

by

Nan Lu

Dissertation submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Horticulture

Richard E. Veilleux, Chair

Vladimir Shulaev

Andy Pereira

Allan W. Dickerman

James G. Tokuhisa

August 30, 2013

Blacksburg, Virginia

Keywords: forward genetics, transformation, transposition, transposable elements, sucrose transporter

Transposon Tagging in Strawberry and Potato and Characterization of Representative Strawberry Mutants

Nan Lu

ABSTRACT

Strawberry and potato are both important crop species in the world providing various nutritional values. The cultivated strawberry, *Fragaria* × *ananassa*, is a fruit crop with a complex genome (2*n*=8*x*=56) whereas the diploid woodland strawberry, *Fragaria vesca*, has a smaller genome (2*n*=2*x*=14, 240 Mb) and lots of other qualities that make it a good model for genetic and genomic study, such as high yield of seeds and efficient transformation. Potato (*Solanum tuberosum*, 2*n*=4*x*=48) is an important vegetable crop in the world and is highly heterozygous. The successful sequencing of the homozygous doubled monoploid clone of potato provides good insight into the study of important genes in this species in improving the pest resistance and improving yield. One approach to characterize gene function in a model system is having large populations of T-DNA insertional or transposon tagged mutants. The idea of using *AcDs* construct to create transposon tagged mutant populations has also been applied in many species. Here we transformed two species, *Fragaria vesca* and a monoploid potato, *Solanum phureja* 1-3-516, which is the progenitor of the sequenced doubled monoploid clone, with the same *AcDs* construct, *Ac-Ds*ATag-Bar_gosGFP, to generate mutant collection, compare the marker gene performance and transposition efficiency, as well as characterizing phenotypic mutants with genes of interest. Transposants were found to reinsert to unlinked sites from the launch pad site in the strawberry genome, whereas in potato transposants tended to locate locally from the launch pad position when using the same construct. One transposon based activation tagging strawberry mutant, with its insertion in the promoter region of gene of interest in strawberry from the *Ac-Ds*ATag-Bar_gosGFP population was studied. In a segregating T_2 population, expression level of the candidate gene, epidermis-specific secreted glycoprotein EP1 precursor, was 670 fold higher in petioles of homozygotes than in wild type plants, suggesting the function of this gene involved in maintaining mechanical strength of petioles. Since the often-used

transposase gene was cloned from the monocot species maize, the efficiency of obtaining germinal transposants was many times lower than expected in order to saturate the genome for diploid species. In order to improve the chance of getting unique transposants, we attempted to codon optimize the transposase gene, as well as switching to microspore specific promoters that had been well characterized to control timing of expression of the transposase gene. Transposants were found in both T_0 primary regenerates and anther culture derived potatoes using both the *pAcDs-AtSCP* and *pAcDs-AmDEFH125* constructs. Sequencing of the empty donor site revealed that excision occurred in different cells during anther culture. A strawberry mutant with sugar transport deficiency due to T-DNA insertion near a sucrose transporter-2 gene showing stunted phenotype with increased level of anthocyanin was also characterized. The concentrations of sucrose, glucose, and fructose were significantly greater in source leaves of the mutant than wild type plants, suggesting these compounds might be substrates of this gene in transporting to sink leaves and roots.

ACKNOWLEDGEMENT

Firstly I would like to thank my advisor, Dr. Richard Veilleux, to give me the chance to perform research here at Virginia Tech, and the opportunity to attend all the national and international conferences. I will never forget his support and encouragement whenever I met any problem in research and in life. I also would like to thank my co-advisor, Dr. Vladimir Shulaev, who gave me a lot of advice on my projects. All other committee members, Dr. Andy Pereira, Dr. Allan Dickerman and Dr. James Tokuhisa, who offered me lot of advice during my committee meeting, my prelim and in research, without which I can not finish all the work.

I would like to thank the Molecular Plant Science program, and especially Dr. Glenda Gillaspy, who offered me the opportunity to come to Virginia Tech five years ago.

My parents, Xuede Lu and Weiyun Yang, supported me all my life and carried me on though all the hard times. And to my girlfriend Yingqi Cai, whose support and help means so much to me.

I would like to thank all my colleagues at Virginia Tech, Dr. Norma Manrique, Dr. Juan Ruiz-Rojas, Dr. Sarah Holt, Kerri Mills, Kendall Upham, Tatiana Boluarte, Brenda Arnold, Hua Xiao, Parker Laimbeer, Sukhwinder Aulakh, for their help, as well as my friends here in Blacksburg, Yihui Fang, Xiaoyan Jia, Will Wadlington, all of whom offered lots of help to me.

Being in graduate school is one of the most unforgettable experiences in my life. I am clearer with my career objectives and I am more confident. I enjoy all the different cultures I experienced. I really appreciate everyone's help here in Virginia Tech.

Thank you,

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION AND LITERATURE OVERVIEW

Introduction of strawberry and potato

Strawberry belongs to the rosaceous family, which consists of over 3,000 species. They include many valuable fruit crops such as apple, peach, pear, blueberry, raspberry, etc. Besides their nutritional value, they also provide important horticultural and economical values to agriculture in temperate regions (Shulaev et al. 2008). With the new technologies available, many species in this family have been or are being sequenced, such as apple (Velasco et al. 2010), peach (International Peach Genome Initiative 2013), strawberry (Shulaev et al. 2011), which provide important information in studying their features. The cultivated strawberry, *Fragaria* × *ananassa*, is an octoploid species (2*n*=8*x*=56) that originated about 300 years ago as a result of crosses between two tetraploid species *Fragaria virginiana* and *Fragaria chiloensis* (Stauct 1988). According to a report from USDA in 2012, the United States is the world's largest strawberry producer, with about 3,000 million pounds in 2012 and constitutes a large proportion in the income. California and Florida are the two most important places for growing strawberries in the United States. An important economical fruit crop as *Fragaria* × *ananassa*, however, its complex genome makes it difficult to study as a model plant species. The woodland strawberry, *Fragaria vesca*, on the other hand, has a simpler and smaller genome (Shulaev et al. 2011), large seed production, as well as efficient transformation protocol (Oosumi et al. 2006). These qualities make it a good model species in this family.

Potato (*Solanum tuberosum*) is one of the most important non-grain food commodities in the world. Belonging to the Solanaceae family, it has lots of unique features compared to other members such as tomato (*Solanum lycopersicum*), tobacco (*Nicotiana tabacum*), or eggplant (*Solanum melongena*) by producing tubers. This cultivated potato species also has a complex genome (2n=4x=48), and is highly heterozygous, which makes it difficult for any genetics study because of the chances of getting inbred lines and getting homozygote mutants are not high. The most recent sequenced homozygote doubled-monoploid clone of potato provides new resources in studying its important features (The Potato Genome Sequencing Consortium 2011).

T-DNA insertion and transposon tagging

There are two most common methods in studying gene functions, forward genetics and reverse genetics. In using forward genetics people generate large populations of insertional mutants hoping to disrupt function of any random genes that could lead to an unusual phenotype; in using reverse genetics method, it starts with some knowledge of a particular gene, and hope to overexpress or knockout this gene of interest and record any data due to the mutation (Alonso and Ecker 2006).

Several labs have succeeded in using forward genetics in the plant community by generating large population of T-DNA insertional mutants in several species such as Arabidopsis (Alonso et al. 2003), rice (Jeon et al. 2000), purple false brome (Vain et al. 2008) and strawberry (Oosumi et al. 2010). However, in order to knockout each gene in a plant genome, it is too labor intensive and depends on transformation efficiency and the chance of getting single insertion mutants.

Transposon tagging has been an alternative to T-DNA insertional mutagenesis ever since the discovery of the jumping gene, transposase, in maize (McClintock 1950). The advantage of using this technique is the possibility of germinal transposition where different genes may be tagged among sibling plants in the next generation or after a few generations. A commonly used strategy for transposon tagging is called the *AcDs* system, which consists of mainly two elements, one *Ac* (*activator*) element and one *Ds* (*dissociator*) element. The *Ac* is autonomous and the *Ds* is non-autonomous, meaning that the function of *Ds* is based on the expression of *Ac*. The *Ds* is usually bordered by IR (inverted repeat) sequences, both of which are about 250 bp long and they are the recognition sites of the enzyme translated from the transposase gene (Mochida and Shinozaki 2010).

Two approaches have been applied in transposon tagging using *AcDs*, either one-component construct, which means both *Ac* and *Ds* occur in the same construct; or two-component construct, in which the *Ac* and *Ds* are in different constructs. The success of using two component constructs is based on the ease of doing cross-pollinations to activate transposition of *Ds* from original insertion site. The chance of getting somatic transpositions at callus or vegetative growth stages is lower using two-component system. Multiple strategies have been developed to optimize the efficiency of using this technique, such as using *cre-lox* system

combined with *AcDs* system (Qu et al. 2009), and using double fluorescent markers for easy selection (Qu et al. 2008).

Successful generation of transposon tagged mutant collections has been accomplished in many plant species using the *AcDs* system, including Arabidopsis (Marsch-MartÍnez 2011; Rosin et al. 2008), poplar (Fladung and Polak 2012), strawberry (Veilleux et al. 2012), rice (Charng 2012; Guiderdoni and Gantet 2012; Jiang et al. 2012; Park et al. 2006), maize (Vollbrecht et al. 2010), barley (Ayliffe and Pryor 2011; Cooper et al. 2004; Zhao et al. 2006), tomato (Carter et al. 2013) and soybean (Mathieu et al. 2009). The problem of using this system, especially in diploid species, is the high frequency of somatic transposition in pre-sporogenous tissue leading to the generation of many seedlings bearing the same transposon tag.

New strategy for improving transposon tagging efficiency

Considerable effort has been made to improve the efficiency of obtaining germinal transposants. The study of p35S promoter fused to the transposase gene in Arabidopsis has proved that transposition tended to occur prior to floral development leading to somatic transposition when using a constitutive promoter (Long et al. 1993). To overcome this obstacle, two-component systems have been applied in many species to prevent transposition before doing crosses. In tomato, over 2,000 stable transposants have been generated (Meissner et al. 2000) using this strategy. Similar results have been reported in rice (Kolesnik et al. 2004; Luan et al. 2008) and broccoli (McKenzie et al. 2002).

Besides using constitutive promoters, other strategies such as using inductive or tissue-specific promoters have also been tried. In a study of rice, using meiosis-associated promoters cloned from lily, the frequency of germinal transposition efficiency was approximately 1% compared to 0.5% using the constitutive promoter p35S (Morita et al. 2003). The development of a system using a heat shock promoter from soybean to control expression of transposase in Arabidopsis has generated a large *Ds*-tagged population localized in a 50 kb region that could be helpful in candidate gene identification (Nishal et al. 2005). This has been important since such results have only been previously reported for a monocot plant species, maize (Moreno et al. 1992).

Sugar transport in higher plants

The ultimate goal of forward genetics is to link phenotypic variation to the genetic background that caused the mutation. Here we picked a strawberry mutant from our population that is related to sugar transport deficiency as an example.

Sugar transport is correlated with photosynthesis in higher plants. Photosynthesis is the process in plants to convert light energy into chemical energy. Light energy and carbon dioxide are converted to carbohydrate and oxygen. This is one of the most important natural ways of producing fixed carbon. And with the increasing need of new clean energy, people are seeking new energy from plants with high yield of biomass, like sugarcane and switchgrass. The high content of sugar in sugarcane makes it ideal to be used to convert to ethanol as new energy, a process that would benefit from a greater understanding of sugar transport in different species.

The successful transportation of carbohydrates that were synthesized in the source leaf tissue to other organs such as roots is fundamental in maintaining basic functions of life. Various sugar transporters exist in plants, such as monosaccharide transporter and hexose transporter (Slewinski 2011). Sugar transporter genes have different affinity for substrates and they have different functions, such as loading or unloading from vascular tissue to sink leaf (Reidel et al. 2009; Turgeon and Wolf 2009), long term transportation in xylem (Johnson et al. 1987; Turgeon 2010), or even in pollen development (Ylstra et al. 1998).

Among these transporters, sucrose transporter is an important member. Based on sequence similarity and functional analysis, the publically available sucrose transporters were divided into four groups (Sauer 2007). The nine well-characterized homologs from Arabidopsis were categorized into three of the four groups, suggesting their different functions within the family.

The study of sucrose transporter genes in fruit crops such as strawberry is important in terms of increasing nutritional values and other qualities. Previous study of *FaSUT1* has demonstrated role of sucrose in strawberry fruit development and ripening (Jia et al. 2013). But the function of other members in this gene family in strawberry has not been characterized as much. Thus the creation of insertional mutant and functional analysis of sucrose transporter gene in strawberry is necessary.

Project Justification and goal

The published draft genome sequences of strawberry and potato provided useful information on candidate genes annotated by bioinformatics tools. To better understand the function of each gene, creation of large insertional or activation tagged mutant populations would be necessary. One good example is the most commonly used model plant species, Arabidopsis. The research on Arabidopsis all over the world is using the seed stock maintained in a few research centers such as the Arabidopsis Biological Resource Center (ABRC). And the Arabidopsis Information Resource website (http://www.arabidopsis.org/) provides detailed information about mutants and the information is available for all research purposes.

For the same purpose, we are building mutant stocks in Virginia Tech and the mutant website (http://hortmutants.vbi.vt.edu/HortMutants/) has been established and maintained.

We used a one-component *AcDs* construct, *Ac-Ds*ATag-Bar_gosGFP, which was synthesized in Dr. Andy Pereira's lab and we transformed it in strawberry and potato to 1) screen for putative transposants, 2) compare the efficiency of marker gene in different plant species, 3) study phenotypic mutants generated from these populations.

For a new *AcDs* construct using microspore-specific promoters, we attempted to increase the efficiency of obtaining germinal transposants by using this new strategy compared to the constructs where transposase is controlled by constitutive promoters, such as in the *Ac-Ds*ATag-Bar_gosGFP construct, or the *pAcDs* construct. The AtSCP (Oh et al. 2010; Oh et al. 2011) and AmDEFH125 (Lauri et al. 2006) promoters were well characterized tissue specific promoters with specific timing of gene expression. We hoped that by using these promoters, we could eliminate most somatic transposants and produce germinal transposants derived after meiosis.

The sugar transport deficient mutant was discovered during screening for transposants of the *AcDs* population. The similarity of mRNA sequences, the high resemblance of mutant phenotype with Arabidopsis and tomato made it a good example of utilizing forward genetics for characterizing and identifying candidate genes.

The research conducted in this project provides a good resource for understanding promoter function in heterologous genomes, for mutant library construction for public use, and characterization of genes by means of forward genetics.

References

Alonso JM, Ecker JR (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. Nat Rev Genet 7:524-536

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Sci Signal 301:653

Ayliffe MA, Pryor AJ (2011) Activation tagging and insertional mutagenesis in barley. Methods Mol Biol 678:107-128

Carter JD, Pereira A, Dickerman AW, Veilleux RE (2013) An active *ac/ds* transposon system for activation tagging in tomato cultivar m82 using clonal propagation. Plant Physiol 162:145-156

Charng YC (2012) A one-time inducible transposon to create knockout mutants in rice. Methods Mol Biol 847:369-377

Cooper LD, Marquez-Cedillo L, Singh J, Sturbaum AK, Zhang S, Edwards V, Johnson K, Kleinhofs A, Rangel S, Carollo V, Bregitzer P, Lemaux PG, Hayes PM (2004) Mapping *Ds* insertions in barley using a sequence-based approach. Mol Genet Genomics 272:181-193

Fladung M, Polak O (2012) *Ac/Ds*-transposon activation tagging in poplar: a powerful tool for gene discovery. BMC Genomics 13:61

Guiderdoni E, Gantet P (2012) *Ac-Ds* solutions for rice insertion mutagenesis. Methods Mol Biol 859:177-187

International Peach Genome Initiative (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet 45:487-494

Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim CJ, Jang S, Lee S, Yang K (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant J 22:561-570

Jia H, Wang Y, Sun M, Li B, Han Y, Zhao Y, Li X, Ding N, Li C, Ji W, Jia W (2013) Sucrose functions as a signal involved in the regulation of strawberry fruit development and ripening. New Phytol 198:453-465

Jiang Y, Cai Z, Xie W, Long T, Yu H, Zhang Q (2012) Rice functional genomics research: progress and implications for crop genetic improvement. Biotechnol Adv 30:1059-1070

Johnson RW, Tyree MT, Dixon MA (1987) A requirement for sucrose in xylem sap flow from dormant maple trees. Plant Physiol 84:495-500

Kolesnik T, Szeverenyi I, Bachmann D, Kumar CS, Jiang S, Ramamoorthy R, Cai M, Ma ZG, Sundaresan V, Ramachandran S (2004) Establishing an efficient *Ac/Ds* tagging system in rice: large-scale analysis of *Ds* flanking sequences. Plant J 37:301-314

Lauri A, Xing S, Heidmann I, Saedler H, Zachgo S (2006) The pollen-specific *DEFH125* promoter from *Antirrhinum* is bound in vivo by the MADS-box proteins DEFICIENS and GLOBOSA. Planta 224:61-71

Long D, Swinburne J, Martin M, Wilson K, Sundberg E, Lee K, Coupland G (1993) Analysis of the frequency of inheritance of transposed *Ds* elements in *Arabidopsis* after activation by a CaMV 35S promoter fusion to the *Ac* transposase gene. Mol Gen Genet 241:627-636

Luan WJ, He CK, Hu GC, Dey M, Fu YP, Si HM, Zhu L, Liu WZ, Duan F, Zhang H, Liu WY, Zhuo RY, Garg A, Wu R, Sun ZX (2008) An efficient field screening procedure for

identifying transposants for constructing an *Ac/Ds*-based insertional-mutant library of rice. Genome 51:41-49

Marsch-MartÍnez N (2011) A transposon-based activation tagging system for gene function discovery in *Arabidopsis*. Methods Mol Biol 754:67-83

Mathieu M, Winters EK, Kong F, Wan J, Wang S, Eckert H, Luth D, Paz M, Donovan C, Zhang Z, Somers D, Wang K, Nguyen H, Shoemaker RC, Stacey G, Clemente T (2009) Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository. Planta 229:279-289

McClintock B (1950) The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A 36:344-355

McKenzie N, Wen LY, Dale J (2002) Tissue-culture enhanced transposition of the maize transposable element *Dissociation* in *Brassica oleracea* var. *′Italica′*. Theor Appl Genet 105:23-33

Meissner R, Chague V, Zhu Q, Emmanuel E, Elkind Y, Levy AA (2000) Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. Plant J 22:265-274

Mochida K, Shinozaki K (2010) Genomics and bioinformatics resources for crop improvement. Plant Cell Physiol 51:497-523

Moreno MA, Chen J, Greenblatt I, Dellaporta SL (1992) Reconstitutional mutagenesis of the maize *P* gene by short-range *Ac* transpositions. Genetics 131:939-956

Morita R, Hattori Y, Yokoi S, Takase H, Minami M, Hiratsuka K, Toriyama K (2003) Assessment of utility of meiosis-associated promoters of lily for induction of germinal *ds* transposition in transgenic rice. Plant Cell Physiol 44:637-642

Nishal B, Tantikanjana T, Sundaresan V (2005) An inducible targeted tagging system for localized saturation mutagenesis in Arabidopsis. Plant Physiol 137:3-12

Oh SA, Park KS, Twell D, Park SK (2010) The *SIDECAR POLLEN* gene encodes a microspore-specific LOB/AS2 domain protein required for the correct timing and orientation of asymmetric cell division. Plant J 64:839-850

Oh SA, Twell D, Park SK (2011) *SIDECAR POLLEN* suggests a plant-specific regulatory network underlying asymmetric microspore division in Arabidopsis. Plant signaling & behavior 6:416-419

Oosumi T, Gruszewski HA, Blischak LA, Baxter AJ, Wadl PA, Shuman JL, Veilleux RE, Shulaev V (2006) High-efficiency transformation of the diploid strawberry (*Fragaria vesca*) for functional genomics. Planta 223:1219-1230

Oosumi T, Ruiz-Rojas JJ, Veilleux RE, Dickerman AW, Shulaev V (2010) Implementing reverse genetics in Rosaceae: analysis of T-DNA flanking sequences of insertional mutant lines in the diploid strawberry, *Fragaria vesca*. Physiol Plant 140:1-9

Park SJ, Piao HL, Xuan YH, Park SH, Je BI, Kim CM, Lee EJ, Park SH, Ryu B, Lee KH, Lee GH, Nam MH, Yeo US, Lee MC, Yun DW, Eun MY, Han CD (2006) Analysis of intragenic *Ds* transpositions and excision events generating novel allelic variation in rice. Mol Cells 21:284-293

Qu S, Desai A, Wing R, Sundaresan V (2008) A versatile transposon-based activation tag vector system for functional genomics in cereals and other monocot plants. Plant Physiol 146:189-199

Qu S, Jeon JS, Ouwerkerk PB, Bellizzi M, Leach J, Ronald P, Wang GL (2009) Construction and application of efficient *Ac-Ds* transposon tagging vectors in rice. J Integr Plant Biol 51:982-992

Reidel EJ, Rennie EA, Amiard V, Cheng L, Turgeon R (2009) Phloem loading strategies in three plant species that transport sugar alcohols. Plant Physiol 149:1601-1608

Rosin FM, Watanabe N, Cacas JL, Kato N, Arroyo JM, Fang Y, May B, Vaughn M, Simorowski J, Ramu U, McCombie RW, Spector DL, Martienssen RA, Lam E (2008) Genome-wide transposon tagging reveals location-dependent effects on transcription and chromatin organization in Arabidopsis. Plant J 55:514-525

Sauer N (2007) Molecular physiology of higher plant sucrose transporters. FEBS Lett 581:2309-2317

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arus P, Dandekar AM, Lewers K, Brown SK, Davis TM, Gardiner SE, Potter D, Veilleux RE (2008) Multiple models for Rosaceae genomics. Plant Physiol 147:985-1003

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton JM, Rees DJ, Williams KP, Holt SH, Ruiz Rojas JJ, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman TL, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Jr., Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Lopez Girona E, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nat Genet 43:109-116

Slewinski TL (2011) Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: a physiological perspective. Mol Plant 4:641-662

Stauct G (1988) The species of Fragaria, their taxonomy and geographical distribution. In: International Strawberry Symposium 265, pp 23-34

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189-195

Turgeon R (2010) The puzzle of phloem pressure. Plant Physiol 154:578-581

Turgeon R, Wolf S (2009) Phloem transport: cellular pathways and molecular trafficking. Annu Rev Plant Biol 60:207-221

Vain P, Worland B, Thole V, McKenzie N, Alves SC, Opanowicz M, Fish LJ, Bevan MW, Snape JW (2008) *Agrobacterium*-mediated transformation of the temperate grass *Brachypodium distachyon* (genotype Bd21) for T-DNA insertional mutagenesis. Plant Biotechnology J 6:236-245

Veilleux RE, Mills KP, Baxter AJ, Upham KT, Ferguson TJ, Holt SH, Lu N, Ruiz-Rojas JJ, Pantazis CJ, Davis CM, Lindsay RC, Powell FL, Dan Y, Dickerman AW, Oosumi T, Shulaev V (2012) Transposon tagging in diploid strawberry. Plant Biotechnol J 10:985-994

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-

Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel CE, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nat Genet 42:833-839

Vollbrecht E, Duvick J, Schares JP, Ahern KR, Deewatthanawong P, Xu L, Conrad LJ, Kikuchi K, Kubinec TA, Hall BD, Weeks R, Unger-Wallace E, Muszynski M, Brendel VP, Brutnell TP (2010) Genome-wide distribution of transposed *Dissociation* elements in maize. Plant Cell 22:1667-1685

Ylstra B, Garrido D, Busscher J, van Tunen AJ (1998) Hexose transport in growing petunia pollen tubes and characterization of a pollen-specific, putative monosaccharide transporter. Plant Physiol 118:297-304

Zhao T, Palotta M, Langridge P, Prasad M, Graner A, Schulze-Lefert P, Koprek T (2006) Mapped *Ds*/T-DNA launch pads for functional genomics in barley. Plant J 47:811-826

2. TRANSPOSON BASED ACTIVATION TAGGING IN DIPLOID STRAWBERRY AND MONOPLOID DERIVATIVES OF POTATO

Abstract

Key message **Diploid strawberry and potato transformed with a transposon-tagging construct exhibited either global (strawberry) or local transposition (potato). An activation tagged compact sized strawberry mutant overexpressed the gene adjacent to *Ds*.**

Abstract Strawberries are important fruit crop species, and potatoes are important vegetable crops, both of which have had their genomes sequenced (Shulaev et al. 2008; The Potato Genome Sequencing Consortium 2011), where 39,031 and 34,809 protein coding genes were predicted in the draft genomes respectively. One effective method to characterize functions of each candidate genes is to have large insertional or activation tagged mutant populations. Here we report the creation of two transposon-tagged mutant populations from these two crops using the same *Activation*/*Dissociation* (*Ac/Ds*) construct. Marker gene expression and transposition ratio showed different results in the two crops. One activation tagged strawberry transposant with a candidate gene, epidermis-specific secreted glycoprotein EP1, was studied. The reasons of different performance of using the same construct in different plant species were also discussed.

Keywords *Solanum tuberosum L.* •, *Fragaria vesca* • transposon tagging, forward genetics, endopolyploidization, epidermis-specific secreted glycoprotein EP1

Introduction

The two crop species, strawberry and potato, are both important sources of various nutrients. Potatoes are highly heterozygous species and are the world's major non-grain food supply (Shulaev et al. 2008; The Potato Genome Sequencing Consortium 2011). A homozygous doubled-monoploid clone of *Solanum tuberosum* Group Phureja used for genome sequencing has a much simpler genome structure compared to cultivated potato and provides a good model for studying candidate genes. Commercial strawberry

cultivars, *Fragaria* × *ananassa*, also have complex genome structure as a highly heterozygous octoploid species. The diploid woodland strawberry, *Fragaria vesca*, on the other hand, provides a useful model for fruit crops in the Rosaceae family. It self-pollinates and tolerates inbreeding and therefore was used to provide a draft genome; it also has a short life cycle, proliferous seed production, transforms readily and has small plant stature permitting the possibility of dense greenhouse plantings (Shulaev et al. 2011). With these emerging model systems in these two vegetatively propagated horticultural crops in addition to their newly available genomic resources, the availability of other genomic toolkits such as large-scale mutant populations would hasten the process of gene discovery.

For the application of forward genetics, one of the most common methods has been to produce large T-DNA insertional mutant populations to select for plants with interesting phenotypes. In the plant kingdom, T-DNA mutant seed stocks have been established in Arabidopsis and rice comprising over 88,000 and 22,090 mutants available for widespread distribution (Alonso et al. 2003; Jeon et al. 2000). A highly efficient transformation protocol, such as *in planta* transformation (Clough and Bent 1998; Hiei et al. 1994), must be available in order to generate such large populations of independent T-DNA insertions. Though both strawberry and potato can be readily transformed, the efficiency is inadequate (Ruiz-Rojas et al. 2010). Transposon tagging, based on the discovery of jumping genes mediated by transposase in maize (McClintock 1950) offers an alternative where it is possible to derive many independent transposon tagged lines from a few independent "launch pad" transformants (Ramachandran and Sundaresan 2001).

The two basic strategies for transposon tagging engage the same components, including transposase, selectable markers, inverted repeats surrounding the *Ds* element, with various promoters used to drive gene expression, but differ in that the *activator* (*Ac*) and *dissociator* (*Ds*) elements occur either in a single T-DNA construct or in two different constructs that are subsequently brought together through cross-pollination. The *Ac* element is comprised mostly of the transposase gene with or without its promoter and a selectable marker whereas the *Ds* element is comprised of two 250 bp IR (inverted repeat) sequences that are recognized by the enzyme expressed from the transposase gene

occurring on either side of a selectable marker. In either case, upon activation, the whole *Ds* element is cut from its original site and reinserted elsewhere in the genome, thereby interrupting or tagging an upstream or downstream native gene. The one component system with *Ac* and *Ds* in the same construct is preferred in species where cross-pollinations are difficult and screening for transposants in the T_1 or subsequent generations is done in self-pollinated progenies. A major difficulty is preventing premature transposition in somatic tissues, a problem which has been addressed by use of inducible or tissue specific promoters (Morita et al. 2003; Nishal et al. 2005) or placing the promoter for transposase in the *Ds* element so that it dissociates from transposase after transposition thereby turning off the process (Veilleux et al. 2012). Germline transposition is ideal as it can result in the generation of unlimited numbers of independent transpositions. The two-component system has *Ac* and *Ds* in separate constructs and crossing is required between an *Ac* transformant and a *Ds* transformant to bring the elements together in a single plant where transposition of *Ds* may occur. Advancing the population to T_2 or further may be required to isolate *Ac* from *Ds* in separate mutant lines and additional transposition may occur until the separation has been effected. Another important factor for an ideal system is the efficiency of marker genes. Either GFP, or antibiotic resistance genes could be used to select for the *Ac* or *Ds* component of the T-DNA. There have been reports of success in generating large transposon-tagged mutant populations in many plant species using either the one-component (Carter et al. 2013; Greco et al. 2003; Veilleux et al. 2012) or two-component (Ayliffe et al. 2007; Brutnell and Conrad 2003; Fladung and Polak 2012; Ito et al. 2005; Kuromori et al. 2004; Mathieu et al. 2009; Vollbrecht et al. 2010) systems. The limitations of our species of interest, i.e., difficult cross-pollination in strawberry and self-incompatibility in diploid potato, have focused our attention on one-component systems.

The *Ac*-*Ds*ATag-Bar_gosGFP construct was first developed for the monocot species, rice (*Oryza sativa*), with monocot promoters driving expression of transposase, Basta herbicide resistance and green fluorescent protein (GFP) genes (Trijatmiko 2005). We have used it with some degree of success in tomato, *Solanum lycopersicum* (Carter et al. 2013). Here we report a comparative study of performance of the same construct, *Ac*-

*Ds*ATag-Bar_gosGFP, in two different horticultural crops with a demonstration of the utility of activation tagging lines with one of the mutants we obtained in strawberry.

Materials and Methods

*Ac-Ds*ATag-Bar_gosGFP Plasmid

The *Ac-Ds*ATag-Bar_gosGFP transposon-tagging construct was kindly provided by Dr. Andy Pereira (Fig. 1). The *Ac* element had green fluorescent protein (GFP) and hygromycin resistance gene (hyg) as selectable markers and the *Ds* element had glufosinate herbicide resistance gene (BAR) as selectable marker. The transposase gene is under control of its native *Ac* promoter. The vector was transformed in *Agrobacterium* strain EHA105 for further plant transformation.

Agrobacterium-mediated plant transformation

For strawberry transformation, wild type *Fragaria vesca* accession number PI551572 was used for transformation according to Oosumi et al. (2006). Up to five regenerated plants from each callus were kept for seed collection. The potato transformation process was modified from An et al. (1986), using the anther-derived monoploid (BARD 1-3 516) of the primitive cultivated potato, *Solanum tuberosum* Group Phureja as starting material. Leaf material from *in vitro* plantlets was used for *Agrobacterium*-mediated transformation. For both experiments, hygromycin was used to select for transformants. GFP expression was also checked under a fluorescent stereomicroscope (Olympus model SZX16®) for transgenic plants.

Flow cytometry

Potato transgenics were processed by flow cytometry to estimate ploidy. Sample preparation was according to Owen et al. (1988). Approximately 0.5 g plant materials were placed in a petri dish and 1 mL of chopping buffer (3.52 mg/ml sodium citrate, 1.7 mg/ml MOPS, 3.7 mg/ml $MgCl_2$, 0.04% v/v Triton X-100) were added. 0.5 mL of filtrate from chopped tissue were pipetted into a microcentrifuge tube. 0.25 mL of ribonuclease (0.8 mg/mL) were then added to each tube and incubated at room temperature for 30 min.

0.125 mL of PI stain (0.4 mg/mL propidium iodide) was added to each sample and incubate on ice for 30 min before running flow cytometry. A monoploid potato was used as control. Samples were run in FACSCalibur flow cytometer (BD Biosciences) using Cellquest Pro acquisition software (BD Biosciences). Results were analyzed using FlowJo software (Treestar, Inc).

Southern blot

Genomic DNA from transgenic strawberry and potato was extracted and digested with *HindIII*. Radioactive probe designed from the GFP gene was used to test copy number of T-DNA insertion with digested DNA (Supplemental Fig. 1).

Self-pollination and cross-pollination for fruit set

Transgenic strawberries were self-pollinated for fruit set and seed collection. Diploid transgenic potatoes identified from flow cytometry were planted in greenhouse and out-crossed as pollen parents with wild type diploid P1 clone as female parent.

Seed germination and growth conditions

Strawberry seeds were sterilized by first washing with 70% ethanol for 5 min, and then kept in sterilized water at 4°C overnight. The next day seeds were soaked in 10% bleach for 15 min, followed by rinsing three times with sterilized water before transferring to liquid germination medium.

For potato seed sterilization, seeds were first washed with 70% ethanol for 5 seconds, and then washed with 30% bleach containing tween 20 for 20 min, followed by rinsing twice with distilled water. Seeds were then soaked in 500 uL of 1000 ppm GA3 overnight. Next day seeds were rinsed with distilled water twice before transferring to liquid germination medium.

Both potato and strawberry seeds were geminated in 125 mL flasks with 10 mL of germination medium (Gamborg's B-5 basal medium, *Phyto*technology Laboratories®, 3% sucrose, PH 5.8) on a shaker (80 rpm) at room temperature. Germinated seedlings were screened for GFP under Olympus SZX16 series microscope.

Strawberry seedlings were grown under controlled environment in a reach-in growth chamber (Conviron, model ATC60) at day neutral condition (12 h photoperiod, 22°C day/16°C night) with light intensity of about 175 μm/m^2/s, before moving to greenhouse conditions. Potato plants were kept under long day condition (16 h photoperiod, 22°C day/18°C night) with the same light intensity.

BASTA® screening and multiplex PCR

Greenhouse plants were screened by painting mature leaves with 0.5% (for strawberry) and 0.03% (for potato) BASTA® and results were recorded after 48 h. Multiplex PCR then was performed to screen for putative transposants. Multiplex PCR was performed using 2x Immomix red polymerase mix (Bioline, BIO-25022) with 1 mM $MgCl_2$, 0.2 mM of each primer (Supplemental Table 3) and 50 ng of DNA in a 25 uL final volume. PCR program for multiplex PCR was one cycle (95°C, 10 min), 35 cycles (94°C for 30 s, 58°C for 30 s, 72°C for 1 min), and a final extension of 5 min at 72°C.

HiTAIL-PCR

A modified HiTAIL-PCR based on Liu and Chen (2007) was used to amplify flanking sequences. Three primer sets (Supplemental Table 3) were used in this study, LBTAIL and *Ds*5-TAIL primers were used to amplify flanking sequences from T_0 launch pads; *Ds*5-TAIL and *Ds*3-TAIL were used to amplify flanking sequences from the inverted repeat part of putative transposants (Supplemental Table 2). Primary or secondary HiTAIL-PCR products were loaded in 1% agarose gel and selected bands were purified according to QIAquick Gel extraction kit for further sequencing. BLAST search was used to anchor sequences in the strawberry genome browser (https://strawberry.plantandfood.co.nz/cgi-bin/nph-blast.cgi?Jform=0) or potato genome browser (http://solanaceae.plantbiology.msu.edu/integrated_searches.shtml) to identify insertion site.

RNA extraction and qRT-PCR

Strawberry RNA was extracted from leaves and petioles according to Yu et al. (2012); 1 ug RNA was treated by *DNAseI* using Ambion DNA-free™Kit (AM1906). cDNA

synthesis and qRT-PCR were performed according to DyNAmo SYBR Green 2-step qRT-PCR Kit. The second step of qRT-PCR was performed in a 20 uL final volume containing 10 uL 2× mastermix, 500 nM of each primer, 0.12 uL of 50× ROX, 100 ng cDNA and added H_2O to final volume in a ABI7500 series real-time PCR system using 96-well optical plates. The program used was followed by the suggested protocol in kit manual. Primer concentration and annealing temperatures were optimized to individual genes. Each experiment was repeated at least three times for analysis, and one blank control with dIH_2O was added. For data analysis we used the relative quantification method.

Results

Transformation

A total of 65 strawberry independent lines was generated after transformation using the *Ac-Ds*ATag-Bar_gosGFP construct (Fig. 1). Among them 12 lines showed early transposition (with *Ac* only or *Ds* only plants) and 53 lines had both *Ac* and *Ds* elements. For potato transformation, no *Ac* or *Ds* only plants were found in any of the 50 T_0 lines tested by multiplex PCR (Table 1).

GFP expression after transformation was strong in strawberry callus (Fig. 2a), regenerating roots (Fig. 2b), flowers (Fig. 2c) and fruits (Fig. 2d). Expression in potato callus was not as strong as in strawberry (Figs. 2e, 2f). It was interesting that in the tomato populations generated, GFP was not strongly expressed either (Carter et al. 2013). Potato and tomato both belong to Solanaceae family while strawberry belongs to Rosaceae family, so the GFP gene with Gos-2 promoter, which was cloned from monocot species behaved differently in different dicot species with a preference to strawberry.

Potato flow cytometry

A total of 150 independent potato transgenic lines was generated after transformation. Among them, 50 lines were selected for flow cytometry where 33 diploid, 12 tetraploid and 5 mixoploid regenerants were found. Due to the genomic complexity of the tetraploid and mixoploid potatoes, only the diploid plants were used for further analysis by crossing with closely related wild type diploid pollinators, P1.

T_1 seedlings screening

Strawberry fruits were collected after self-pollination. Germinated seedlings were screened for GFP (Fig. 3a) and herbicide resistance (Fig. 3b). For strawberry T_1 screening, 29 families showed 3:1 GFP+:wild type ratio, suggesting a single T-DNA insertion. Seven families had a GFP+:wild type ratio of 15:1, suggesting double T-DNA insertions. Other families appeared to have multiple insertions. For these 36 families, only GFP negative plants were kept for further analysis. 0.5% BASTA® was used to paint the mature leaves of plants in the greenhouse and results were recorded after 48 h. BASTA® resistant plants were kept for DNA analysis. Herbicide resistance testing was 80% accurate compared with PCR results in families tested (data not shown).

For potato, no obvious GFP was observed in any of the progenies tested (Fig. 3c), similar to the tomato lines that carried the same GFP construct (Carter et al. 2013). BAR selection was effective in screening for putative transposants (Fig. 3d), with 85% accurate with PCR results (data not shown).

For potato seedlings, in 25 families, the expected 1 transgenic:1 wild type segregation was observed, indicating a single gene insertion without transposition of *Ds*. Six other families segregated 3 transgenic:1 wild type, indicating two insertions without transposition. Unexpected segregation patterns were found among 200 progeny of seven families from three independent T_0 plants, 50% were *Ds* only plants, and 50% were *AcDs* plants without any wild types found. One theory to explain this was the launch pad had early transposition in a diploid cell and went through mitotic recombination where the resulting two somatic cells carried *Ac-Ds*//*Ds* and *Ac*//wt. Then fruits developed from *Ac-Ds*//*Ds* cells would have *Ac-Ds*//wt and *Ds*//wt progeny.

Since the starting material was leaf tissue from monoploid potato, then the question became when and which cell types were the targets of T-DNA insertion during transformation. If monoploid cells were the targets, then the regenerated potato lines would be homozygous for T-DNA and no wild type plants were expected after outcrosses. If the diploid cells were the targets, then different ratios of transgenic:wild type would be possible. These results confirmed that the diploid cells after

endopolyploidization was the preferable target of T-DNA insertion for *Agrobacterium*-mediated transformation.

Multiplex PCR

Multiplex PCR was used for further confirmation of putative transposants. Primers used to amplify housekeeping gene, alcohol acyltransferase (*FvAAT3*, Genbank accession number AF193790) for strawberry and solanidine galactosyltransferase 1 (*SGT1*, Genebank accession number U82367) for potato, were used as control primers for satisfactory DNA extraction; hygromycin and BAR primers were used to amplify segments of the *Ac* and *Ds* elements, respectively. Samples that exhibited only BAR and housekeeping gene bands were considered putative transposants (Figs. 4a, b). In screening among different families, the ratio of families with transposition varied, with 31% in strawberry and 16% in potato (Table 2).

Insertion sites in genome

T_0 launch pad insertion sites were located by HiTAIL-PCR from either left border or right border. Flanking sequences from putative transposant insertion sites were amplified using *Ds*TAIL primers either from 5' or 3' end in IR.

For strawberry characterization, 35 unique insertion sites of launch pad sequences were identified, together with 28 unique transposant insertion sites identified from 187 candidate transposants in PCR screening of 1,332 T_1 plants (Fig. 5a). Twenty-nine unique potato transposants were found from three active launch pads in 90 candidate transposants out of 622 plants, among which 26 were from the same potato T_0 launch pad, 36B. (Fig. 5b)

The *Ds* insertion sites in the three diploid genomes, strawberry, potato and tomato, all had the highest ratio in genes (exon or intron), a relatively similar small ratio in 3'UTR region (1-300 bp downstream of stop codon). The ratios of insertion in promoter region and intergenic region had larger variation (7%-29% for promoter region, 4%-47% for intergenic region, Table 3). These data were consistent with the published strawberry transposon tagging population, where they had similar ratio of insertion in promoter

region, while they also have highest ratio in gene region, and variable in 3'UTR and intergenic regions.

Based on these results, the distribution of *Ds* insertion in all four populations all had a preference toward genes, with low ratio in 3'UTR region, and variable in promoter or intergenic regions. Similar results have been found in other reports (Schneidereit et al. 2008).

Localized transposition of potato line 36B

In the study of generating transposon tagged mutants in Arabidopsis using *AcDs* construct, lots of local transposition of *Ds* were found near a small region in a certain chromosome. For one of the potato lines that we analyzed, 36B, such phenomenon was observed. The original launch pad was found in chromosome 11 within gene PGSC0003DMT400047772, and the screening of T_1 plants derived from this launch pad ended up with 26 unique transposants, most of which were in the same region (Fig. 5b). Among the 26 transposants, 8 of them were near or within a 90 kb distance, where 13 candidate genes were annotated. Interestingly, 9 of the 13 genes were disease resistance related (Fig. 6). This would provide a good resource for cloning and studying these disease resistance genes.

Abnormal excisions

Two somaclones from each T_0 strawberry line carrying either *Ac* only or *Ac/Ds* elements were tested to amplify empty donor sites (EDS) to check existence of footprint. Amplification was successful in seven families, while for one family, SH38, amplification was found in only one somaclone but not the other. Two abnormal excision events were found after sequencing EDS PCR products (Fig. 7). In both cases, *Ds*-5'IR part was still intact with *Ac* element while 3'-IR-enhancer-ubi-BAR part was excised. Abnormal excision could be useful for studying chromosome rearrangement (Zhang et al. 2011).

Zygosity test

Zygosity test was only performed in strawberry *Ds* plants because of the self-incompatibility of potato lines. Seeds from four T_1 strawberry transposant families were selected for zygosity tests in the T_2. Homozygous strawberry plants were found in three families, T_1SH47E-7, T_1SH5D-3 and SH23C-18 (Table 4, Fig. 8a).

T_2 mutant analysis

For one of the T_2 segregating transposant families, T_1SH47E-7, a phenotypic difference was observed between wild type and homozygous mutants. The homozygous plants were more compact compared to wild type plants (Fig 8b). Other morphological traits such as plant height, fruit number and size were not significantly different (data not shown). Unlike many other species in Rosaceae, strawberries have short stems (rosettes) and long petioles.

The insertion of the *Ds* element in T_1SH47E-7 was 579 bp upstream of the start codon of hybrid gene model gene27253 in the strawberry genome browser, where the 4xp35S enhancer was in the same direction as the gene. In a 20 kb region from the insertion site, a cluster of epidermis-specific secreted glycoprotein EP1 gene family members were found (gene27251, gene27253, gene27254, gene27255), as well as a Mediator of RNA polymerase II transcription subunit 1.1 (gene27252, Fig. 8c).

In order to test expression levels of these four genes due to *Ds* insertion, four homozygous, four hemizygous and four wild type T_2 plants were processed for semi-quantitative RT-PCR analysis. No amplification was found for gene27251 in any plant except DNA control, indicating it might be a pseudogene or not expressed in leaves. Similar amplifications were found for gene27252 and gene27254 among the wild type, hemizygous and homozygous plants. However, for gene27253 which was most adjacent to the activation tag insertion, expression was enhanced in homozygous and hemizygous plants compared with wild type plants in both petiole and leaf tissues (Fig. 8d).

In order to better quantify the differences, a real time qRT-PCR experiment was further conducted to quantify the expression level change of gene27253. In petioles the expression was 670-fold higher in homozygotes compared with wild type plants; in leaf

tissues the expression was 298-fold greater. This result was consistent with the band intensity on the gel (Figs. 8d, 8e). Thus it would be reasonable to assume the importance of gene27253 in maintaining structure and mechanical strength in petioles and leaves of *Fragaria vesca*. The ortholog of gene27253 in Arabidopsis is AT1G78860, a curculin-like (mannose-binding) lectin family protein. The expression of AT1G78860 in eFP browser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) showed greatest expression in petioles and leaves (Fig. 8f). This is consistent with the assumption that this gene is related to mechanic strength of the cell wall. The amino acid sequence of gene27253, which belongs to the β-lectin superfamily, was then aligned against the NCBI protein database using neighbor-joining tree method, where max sequence difference was set at 0.50, and many proteins with highly significant e-values were found on the phylogenic tree, meaning there should be conserved regions in this gene family (Fig. 8g).

Mutant collection

Besides the T_2 mutant mentioned above, a few other mutants with distinct phenotypes were also found during our study. They include dwarf (Fig. 9a), yellow leaves (Fig. 9b), compact and high ploidy (Fig. 9c), small leaves and small fruits (Fig. 9d). These mutant phenotypes did not segregate with the T-DNA insertions.

Discussion

Establishment of T-DNA or transposon tagged mutants

The establishment of insertional mutant libraries has greatly contributed to the genomic research in plant and animal kingdoms. The effort of optimizing *AcDs* constructs is continuing. And to different species different systems were chosen due to the self-compatibility. The use of two-component system tends to reduce somatic transposants because transposition will not occur before crosses are done. But it might be necessary to go through another two or more generations to segregate *Ds* from *Ac* and derive homozygous *Ds* mutants. The one-component system could reduce much labor, especially for plants like strawberry where cross-pollinations are difficult. In our report we created two populations of transposon-tagged mutants in two horticultural crop species using the same *AcDs* construct, and compared transposition efficiency.

Compared to the *Ac/Ds* construct in which the promoter controlling transposase was integrated in the *Ds* element, the *Ac-Ds*ATag-Bar_gosGFP construct had the *Ac* promoter linked directly with the transposase gene in the *Ac* element. Hence, transposition can occur repeatedly until EDS and *Ds* elements segregate. Strawberries had more families with transposants but many were somatic. Potato had one family with a common insertion and another family with lots of local transposition. However the transposition ratio of these two crops are much lower than the published tomato activation tagged lines using the same construct (Carter et al. 2013).

The 28 strawberry transposants identified were from multiple families while the 29 potato transposants were basically from the same launch pad. The strawberry transposants were not linked to their launch pad (global transposition) while potato transposants preferred local transposition linked to the launch pad.

Marker gene efficiency in different species

Our finding provides new evidence how promoter activity changes in different species. Expression of the marker gene, GFP, which is under the control of the Gos-2 promoter cloned from maize, behaves totally differently in three dicot species, strawberry, potato and tomato (Carter et al. 2013). Previous study of the Gos-2 promoter shows that it is more active in monocots than in dicots (Assem et al.). The anther specific promoters cloned from rice lost its specificity in dicot Arabidopsis, behaving as constitutive promoters instead (Khurana et al. 2013). The transposase gene in *Ac-Ds*ATag-Bar_gosGFP construct is also under control of the native *Ac* promoter from maize. In our research we found somatic transposition in strawberry primary regenerated plants, but not in potatoes. This could be due to the long time period of transformation of strawberry (10-12 months) compared to potato (2-3 months). Transgene expression in a host genome has been suggested to relate to copy number of T-DNA insertions but not with insertion site, and the high level of transgene expression could trigger RNA silencing (Hamilton and Baulcombe 1999; Schubert et al. 2004). It was also proposed that transgene expression corresponds to gene methylation, which increased as a result of high copy number of T-DNA insertion, including transposase gene (Eun et al. 2012). The reason of higher somatic transposition in strawberry but not in potato could be because of the long

dynamic environment in tissue culture that triggered transposase gene expression. In a previous report, it was found that dicot promoters cloned from one dicot species, *Medicago truncatula,* have different expression activity in different dicot species (Untergasser et al. 2012).

Diploid somatic cells are target for T-DNA insertion in potato transformation

The starting material for potato transformation, the monoploid potato *Solanum phureja* 1-3-516, which was developed by anther culture, is the progenitor of the double monoploid (DM) used for the draft potato genome sequence. The regenerated diploid and tetraploid lines were results of endopolyploidization (Fig. 10). The *Ac*:*Ds* ratio in the T_1 potato progeny provided evidence that diploid cells occurring in the monoploid leaves were the primary targets of T-DNA insertion during transformation since homozygous T-DNA would not produce *Ac* or *Ds* only plants in T_1 progeny (Fig. 10).

New technologies and challenges

New technologies with the possibility of knockout genes of interest are emerging such as TALEN (Cermak et al. 2011) and CRISPR/Cas (Cong et al. 2013). These next-generation genome-editing technologies use endonuclease containing specific sequences targeting genes of interest will be useful alternative strategies. Other strategies such as RNAi (Fire et al. 1998) or ZFN (Urnov et al. 2010) are also needed. However, a mutant collection for plants like strawberry, which requires long time period of transformation, is still necessary. Besides, transposon tagging is a versatile strategy that the transposants could be either knockout, knockdown or overexpression mutants, as shown in our report.

All transposants collected will be integrated into an online website (http://hortmutants.vbi.vt.edu/HortMutants/).

All sequencing results from HiTAIL-PCR and primers used in this report were in Supplemental Material (Supplemental Tables 1, 2, 3).

Acknowledgments

This work was supported by USDA⁄ NRI Award No. 2008-35300-04458, USDA⁄ NIFA Award Nos. VAW-2010-01562 and VAW-2009-04069, USDA Award 10-2361-04 (Biodesign and Bioprocessing Center), The Virginia Tobacco Indemnification Program, Virginia Agricultural Experiment Station Hatch Project 135816, the Virginia Tech Department of Horticulture and the Virginia Tech Molecular Plant Sciences Program.

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301:653-657

An G, Watson BD, Chiang CC (1986) Transformation of Tobacco, Tomato, Potato, and *Arabidopsis thaliana* Using a Binary Ti Vector System. Plant Physiol 81:301-305

Assem SK, El-Itriby HA, Saad M, Madkour MA (2001) Comparison of the efficiency of some novel maize promoters in monocot and dicot plants. Arab J Biotech 5:57-66

Ayliffe MA, Pallotta M, Langridge P, Pryor AJ (2007) A barley activation tagging system. Plant Mol Biol 64:329-347

Brutnell TP, Conrad LJ (2003) Transposon tagging using *Activator (Ac)* in maize. Methods Mol Biol 236:157-176

Carter JD, Pereira A, Dickerman AW, Veilleux RE (2013) An active *ac/ds* transposon system for activation tagging in tomato cultivar m82 using clonal propagation. Plant Physiol 162:145-156

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16:735-743

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819-823

Eun CH, Takagi K, Park KI, Maekawa M, Iida S, Tsugane K (2012) Activation and epigenetic regulation of DNA transposon *nDart1* in rice. Plant Cell Physiol 53:857-868

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806-811

Fladung M, Polak O (2012) *Ac/Ds*-transposon activation tagging in poplar: a powerful tool for gene discovery. BMC Genomics 13:61

Greco R, Ouwerkerk PB, De Kam RJ, Sallaud C, Favalli C, Colombo L, Guiderdoni E, Meijer AH, Hoge Dagger JH, Pereira A (2003) Transpositional behaviour of an *Ac/Ds* system for reverse genetics in rice. Theor Appl Genet 108:10-24

Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286:950-952

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J 6:271-282

Ito T, Motohashi R, Kuromori T, Noutoshi Y, Seki M, Kamiya A, Mizukado S, Sakurai T, Shinozaki K (2005) A resource of 5,814 *dissociation* transposon-tagged and sequence-indexed lines of *Arabidopsis* transposed from start loci on chromosome 5. Plant Cell Physiol 46:1149-1153

Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Lee S, Yang K (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant J 22:561-570

Khurana R, Kapoor S, Tyagi AK (2013) Spatial and temporal activity of upstream regulatory regions of rice anther-specific genes in transgenic rice and *Arabidopsis*. Transgenic Res 22:31-46

Kuromori T, Hirayama T, Kiyosue Y, Takabe H, Mizukado S, Sakurai T, Akiyama K, Kamiya A, Ito T, Shinozaki K (2004) A collection of 11 800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. Plant J 37:897-905

Liu YG, Chen Y (2007) High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. BioTechniques 43:649-650, 652, 654 passim

Mathieu M, Winters EK, Kong F, Wan J, Wang S, Eckert H, Luth D, Paz M, Donovan C, Zhang Z, Somers D, Wang K, Nguyen H, Shoemaker RC, Stacey G, Clemente T (2009) Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository. Planta 229:279-289

McClintock B (1950) The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A 36:344-355

Morita R, Hattori Y, Yokoi S, Takase H, Minami M, Hiratsuka K, Toriyama K (2003) Assessment of utility of meiosis-associated promoters of lily for induction of germinal *ds* transposition in transgenic rice. Plant Cell Physiol 44:637-642

Nishal B, Tantikanjana T, Sundaresan V (2005) An inducible targeted tagging system for localized saturation mutagenesis in Arabidopsis. Plant Physiol 137:3-12

Oosumi T, Gruszewski HA, Blischak LA, Baxter AJ, Wadl PA, Shuman JL, Veilleux RE, Shulaev V (2006) High-efficiency transformation of the diploid strawberry (*Fragaria vesca*) for functional genomics. Planta 223:1219-1230

Owen HR, Veilleux R, Haynes F, Haynes K (1988) Variability for critical photoperiod for tuberization and tuber yield among monoploid, anther-derived genotypes of *Solanum phureja*. J Amer Soc Hort Sci 113:755-759

Ramachandran S, Sundaresan V (2001) Transposons as tools for functional genomics. Plant Physiol Biochem 39:243-252

Ruiz-Rojas JJ, Sargent DJ, Shulaev V, Dickerman AW, Pattison J, Holt SH, Ciordia A, Veilleux RE (2010) SNP discovery and genetic mapping of T-DNA insertional mutants in *Fragaria vesca* L. Theor Appl Genet 121:449-463

Schneidereit A, Imlau A, Sauer N (2008) Conserved *cis*-regulatory elements for DNA-binding-with-one-finger and homeo-domain-leucine-zipper transcription factors regulate companion cell-specific expression of the *Arabidopsis thaliana SUCROSE TRANSPORTER 2* gene. Planta 228:651-662

Schubert D, Lechtenberg B, Forsbach A, Gils M, Bahadur S, Schmidt R (2004) Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. Plant Cell 16:2561-2572

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arus P, Dandekar AM, Lewers K, Brown SK, Davis TM, Gardiner SE, Potter D, Veilleux RE (2008) Multiple models for Rosaceae genomics. Plant Physiol 147:985-1003

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton JM, Rees DJ, Williams KP, Holt SH, Ruiz Rojas JJ, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman TL, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Jr., Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Lopez Girona E, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nat Genet 43:109-116

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189-195

Trijatmiko KR (2005) Comparative Analysis of Drought Resistance Genes in Arabidopsis and Rice. Wageningen Universiteit

Untergasser A, Bijl GJ, Liu W, Bisseling T, Schaart JG, Geurts R (2012) One-step Agrobacterium mediated transformation of eight genes essential for rhizobium symbiotic signaling using the novel binary vector system pHUGE. PLoS One 7:e47885

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11:636-646

Veilleux RE, Mills KP, Baxter AJ, Upham KT, Ferguson TJ, Holt SH, Lu N, Ruiz-Rojas JJ, Pantazis CJ, Davis CM, Lindsay RC, Powell FL, Dan Y, Dickerman AW, Oosumi T, Shulaev V (2012) Transposon tagging in diploid strawberry. Plant Biotechnol J 10:985-994

Vollbrecht E, Duvick J, Schares JP, Ahern KR, Deewatthanawong P, Xu L, Conrad LJ, Kikuchi K, Kubinec TA, Hall BD, Weeks R, Unger-Wallace E, Muszynski M, Brendel VP, Brutnell TP (2010) Genome-wide distribution of transposed *Dissociation* elements in maize. Plant Cell 22:1667-1685

Yu D, Tang H, Zhang Y, Du Z, Yu H, Chen Q (2012) Comparison and improvement of different methods of RNA isolation from strawberry (*Fragria* x *ananassa*). J Agric Sci 4:51-56

Zhang J, Yu C, Krishnaswamy L, Peterson T (2011) Transposable elements as catalysts for chromosome rearrangements. Methods Mol Biol 701:315-326

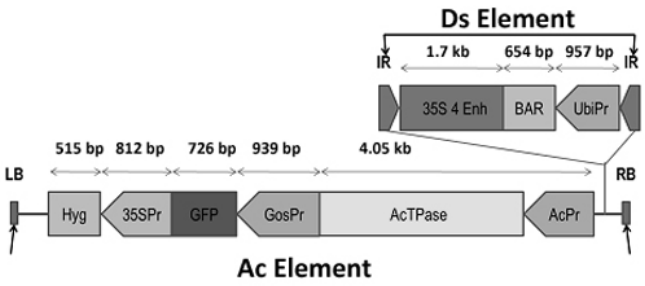


Figure 1 The *Ac*-*Ds*ATag-Bar_gosGFP construct. Elements are as follows; LB = Left Border, Hyg = hygromycin resistance, 35SPr = Cauliflower Mosaic Virus 35*S* promoter, GFP = Fluorescent Protein (Genbank Accession ABB59985), GosPr = *Zea mays* Gos-2 promoter, AcTPase = *Zea mays* transposase (Genbank Accession X05424), AcPr = *Zea mays* transposase native promoter, IR = Inverted Repeat, 35S 4 Enh = tetramer of Cauliflower Mosaic Virus 35S enhancer, BAR = glufosinate resistance, UbiPr = Ubiquitin promoter, IR = Inverted Repeat, RB = Right Border.

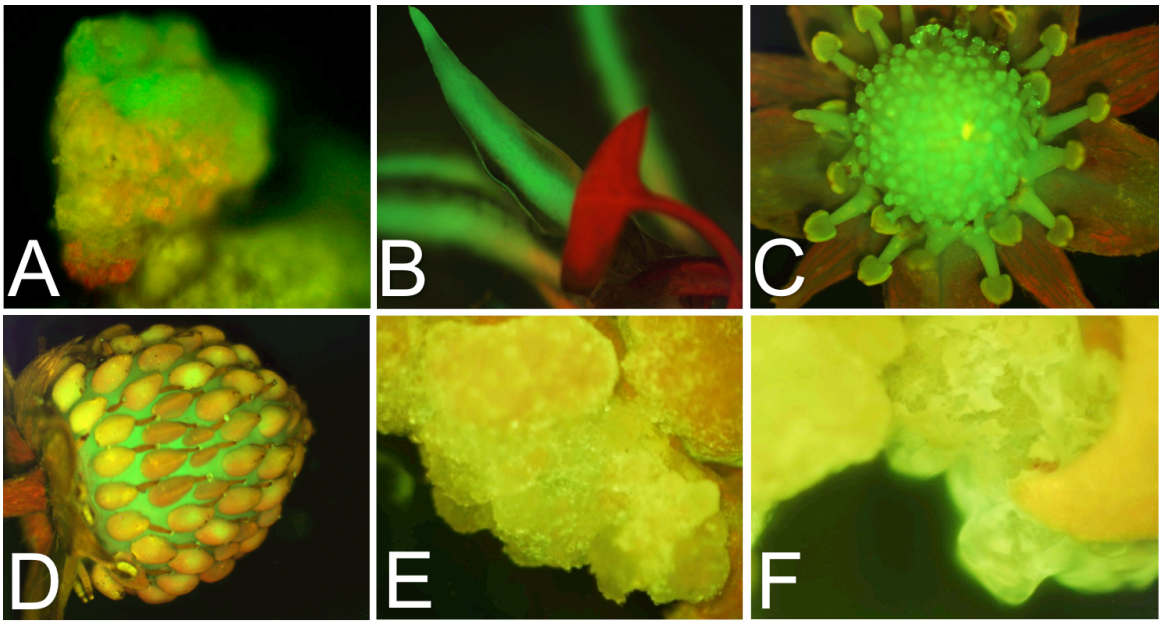


Figure 2 GFP expression in transgenic strawberry and potato after transformation with *Ac-Ds*ATag-Bar_gosGFP construct. Strong expression in **a** strawberry callus, **b** root, **c** flower and **d** fruit, as well as weak GFP expression in potato calli (**e**, **f**).

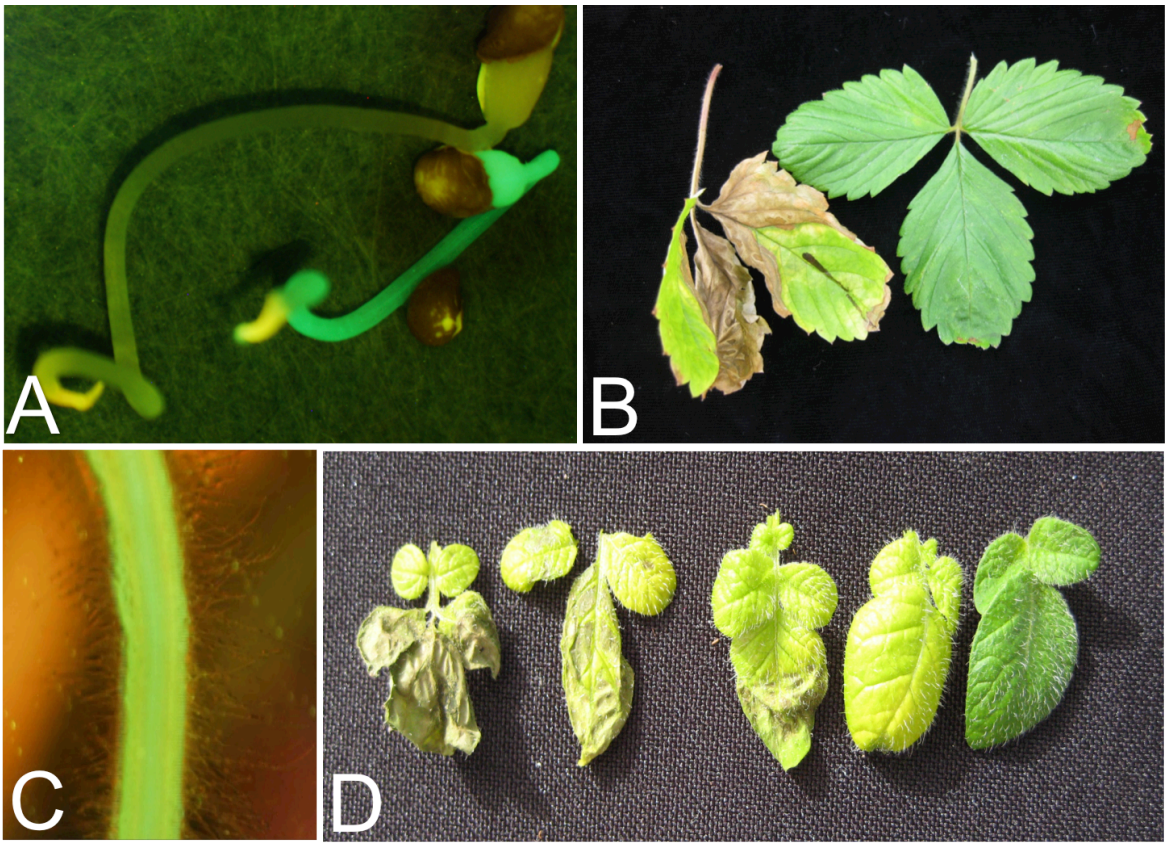


Figure 3 GFP and Basta screening for transposants in T_1 plants derived from self-pollination (strawberry) or cross-pollination (potato): **a** Strawberry GFP negative (left) and positive (right) seedlings 2 weeks after germination. **b** Strawberry Basta sensitive (left) and resistant (right) mature leaves 48 h after leaf painting. **c** GFP expression in potato vascular tissues. **d** Potato Basta sensitive (left 1-4) and resistant (rightmost) leaves.

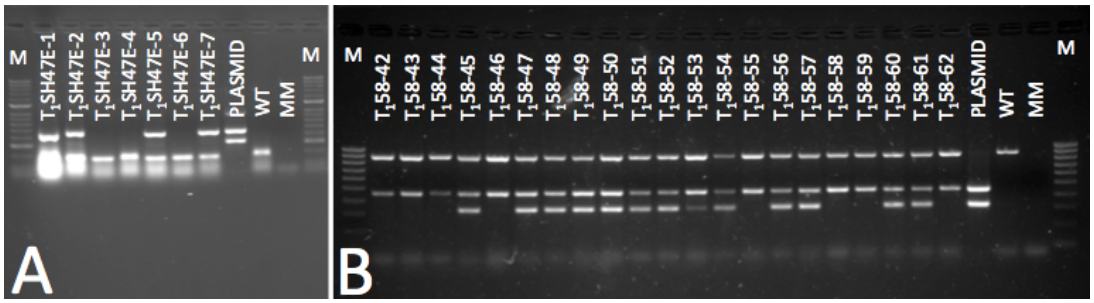
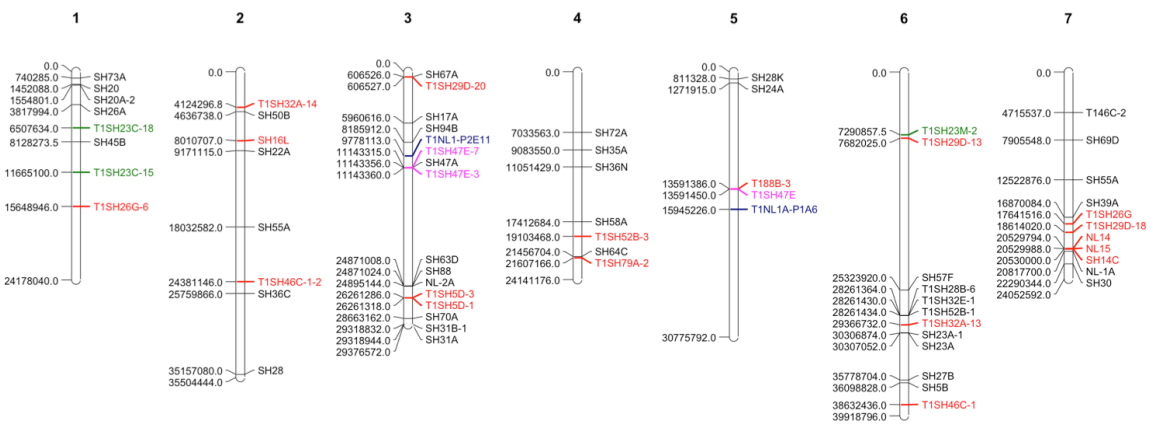


Figure 4 Multiplex PCR used for screening transposants in T_1 strawberry and potato plants. Primers designed for housekeeping gene (*FvAAT3* for strawberry, 175 bp; *SGT1* for potato, 900 bp), Hygromycin (300 bp), BAR (450 bp) were mixed in a single PCR reaction to select for *Ds* only plants. **a** Screening of strawberry transposants. T_1SH47E1-7 (lanes 2-8) were strawberry plants, Plasmid (lane 9), wild type (WT, lane 10) and PCR mastermix served as positive and negative controls. Lanes 1 and 12 were HypperLadder™ II. T_1SH47E-2, 5, 7 were putative transposants. **b** Screening of potato transposants. $T_1$58-42 – 62 (lanes 2-22) were potato plants, plasmid (lane 23), DM wild type potato (WT, lane 24) and PCR mastermix (lane 25) served as positive and negative controls. Lane 1 and 26 were HyperLadder™ IV. $T_1$58-42, 43, 44, 46, 55, 58, 59, 62 were putative transposants.

A)



B)

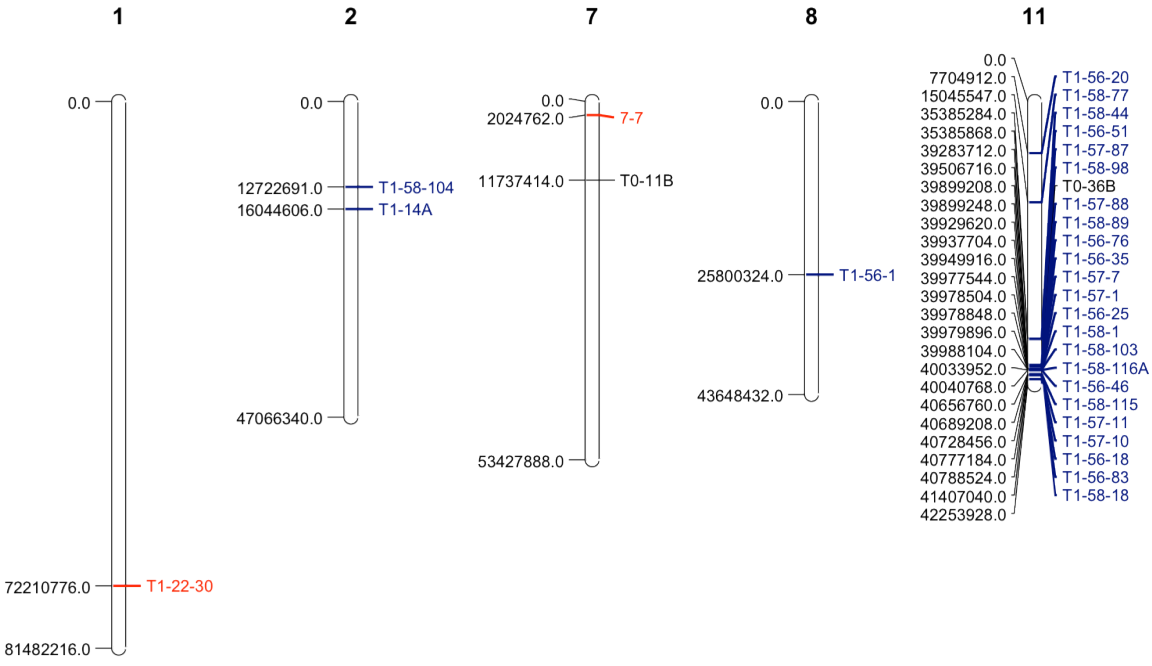


Figure 5 Insertion sites of **a** strawberry and **b** potato mutants on physical maps. Numbers on top of the ideograms refer to the pseudochromosomes of strawberry (**a**) or potato (**b**) where transformants or transposants were situated. Numbers on the left of each pseudochromosome show the physical position (in bp) of each mutant. Names on the right of pseudochromosome show ID of each mutant. Black IDs show T_0 launch pads. Blue, green, purple IDs in (**a**) show unique transposants from three different families in strawberry; blue IDs in (**b**) show unique transposants from one family in potato. Red IDs in (**a**) and (**b**) show unique transposants from various families in strawberry or potato. Only potato pseudochromosomes with T-DNA insertions are shown in **b**. Program used was Mapchart 2.2 software.

Table 1 Number of putative launch pads and early transposed T_0 plants after transformation in strawberry and potato.

Category	Strawberry transformation	Potato transformation
AcDs lines	53 (81.5%)	50 (100%)
Ac or *Ds* only lines	12 (18.5%)	0 (0%)
Total	65 (100%)	50 (100%)

Table 2 Comparison of number of families showing transposition among strawberry, potato and tomato.

Category	Strawberry	Potato	Tomato (a)
No. of families with transposition	20 (31%)	8 (16%)	13 (52%)
No. of families without transposition	45 (69%)	42 (84%)	12 (48%)
Total	65	50	25

(**a**) Data is from (Carter et al. 2013)

Table 3 Categorized T-DNA or *Ds* insertion sites in respective genomes.

Species	Promoter+5'UTR	Exon+Intron	3'UTR	Intergenic
T_0 strawberry	6 (16%)	14 (38%)	1 (3%)	16 (43%)
T_0 potato	0	1	0	1
T_1 strawberry	8 (29%)	17 (61%)	2 (7%)	1 (4%)
T_1 potato	3 (7%)	15 (52%)	2 (7%)	9 (47%)
T_1 Tomato (a)	15%	43%	6%	36%
T_1 *AcDs* strawberry (b)	30%	38%	17%	15%

(**a**) Data from (Carter et al. 2013)
(**b**) Data from (Veilleux et al. 2012)

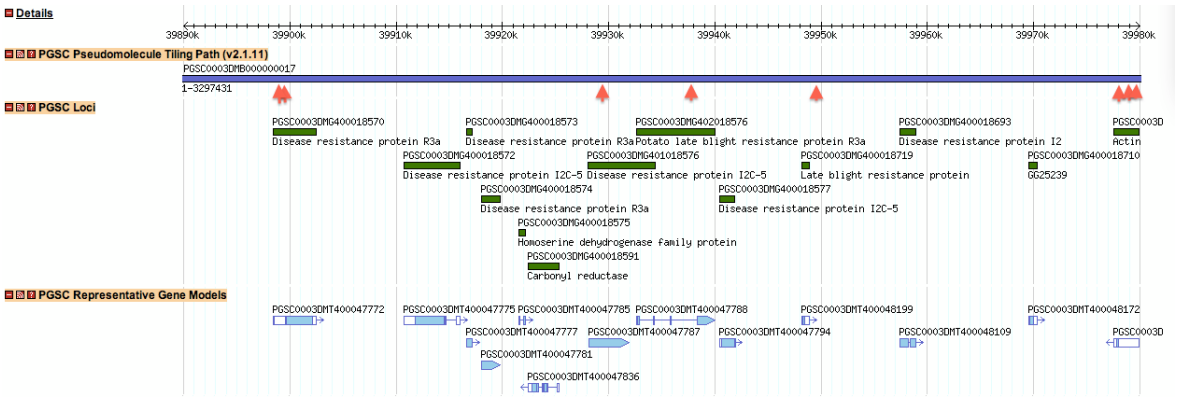


Figure 6 Insertion sites of unique potato transposants developed from T_0 launch pad 36B in a 90 kb region on potato chromosome 11. Nine of the 13 candidate genes within this region annotated from PGSC model were disease resistance related. The eight orange arrows show *Ds* insertion sites within candidate disease resistance related genes. Launch pad 36B insertion site is near the first orange arrow on left.

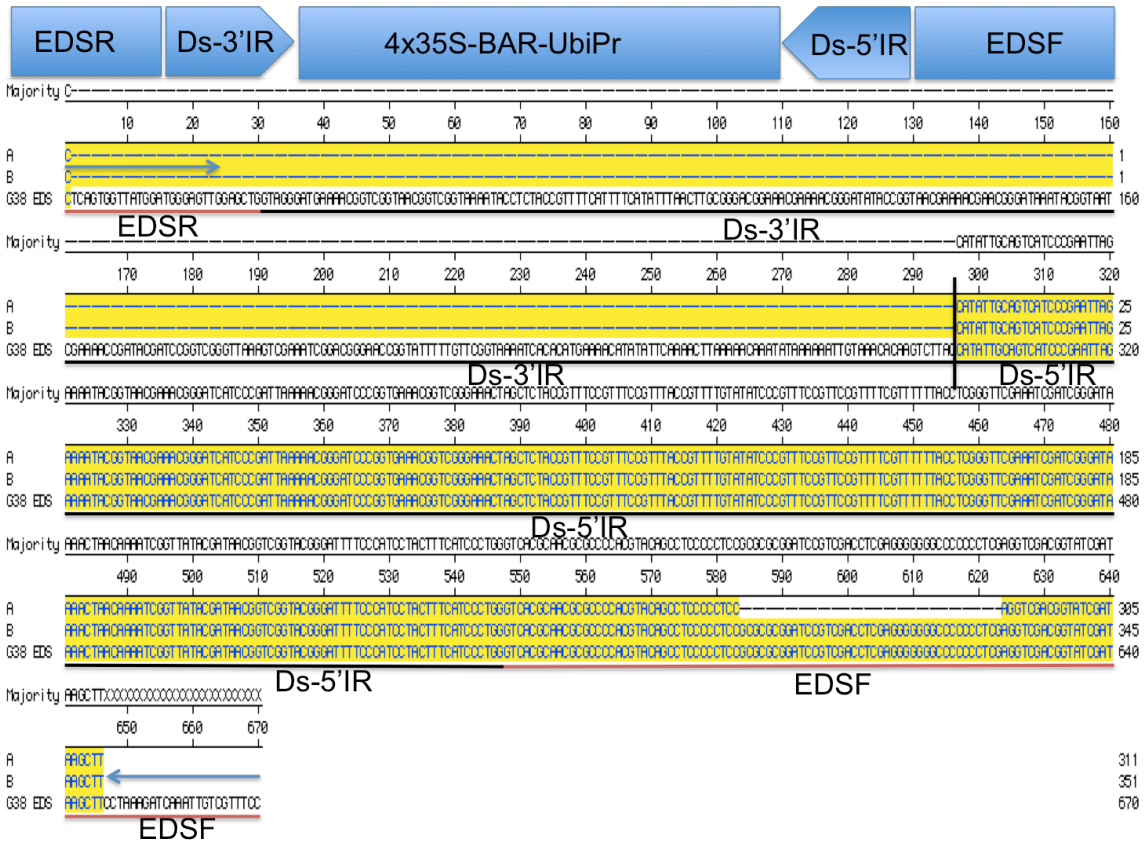
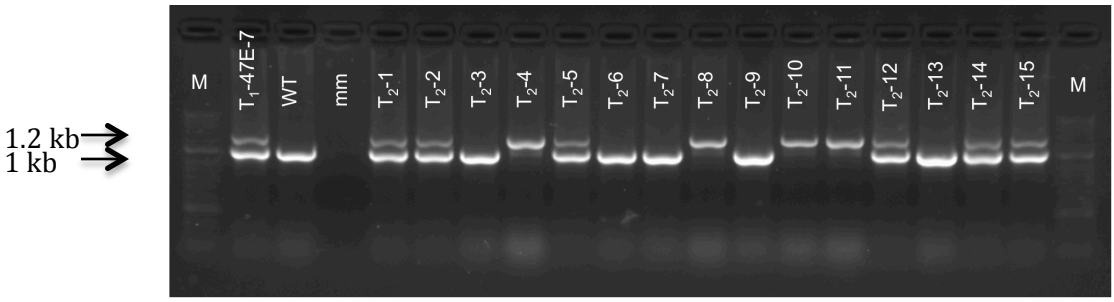


Figure 7 Abnormal excision events in strawberry mutants. A and B were two excision patterns from EDS PCR sequencing results after amplification with G38EDSF and G38EDSR primers. The 4x35S-BAR-UbiPr sequence between *Ds*-3'IR and *Ds*-5'IR is not shown in picture but occurs where the black line at 297 bp is indicated. A was missing a 40 nt fragment in EDSF part. Both A and B had *Ds*-5'IR still intact with *Ac* after *Ds* excision, and *Ds*-3'IR was excised together with *Ds*. EDSR and EDSF mean the region where primers were designed from and blue arrows indicate G38EDSF and G38EDSR primers.

Table 4 Zygosity testing of T_2 seedlings from four strawberry *Ds* families. In each family, 10-20 T_2 plants were screened using one T-DNA specific primer and two gene specific primers to identify homozygous, wild type and hemizygous plants. Chi-square test was used to test the hypothesis of *Ds* being single T-DNA insertion.

ID	Homozygote	Hemizygote	Wild type	χ^2
T_1SH47E-7	4	6	5	0.73
T_1SH5D-3	3	5	3	0.09
SH23C-18	10	5	5	7.5
SH88B-3	18		9	1

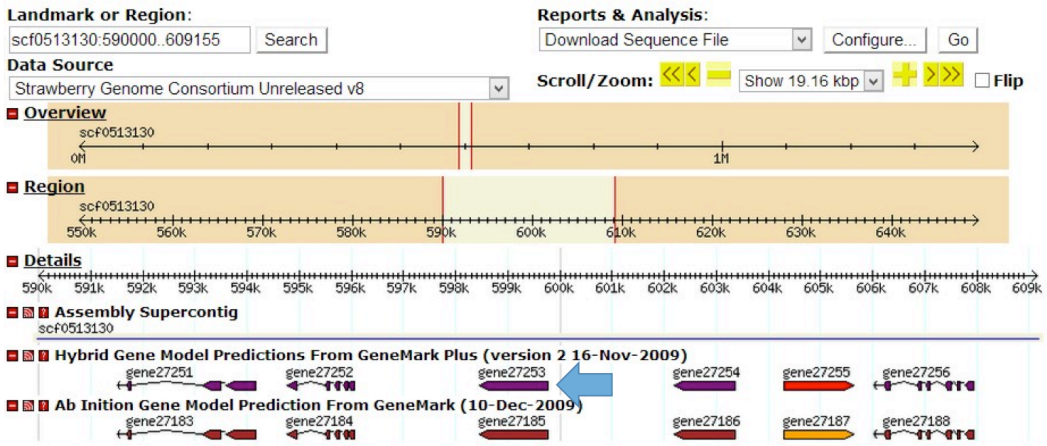
A)



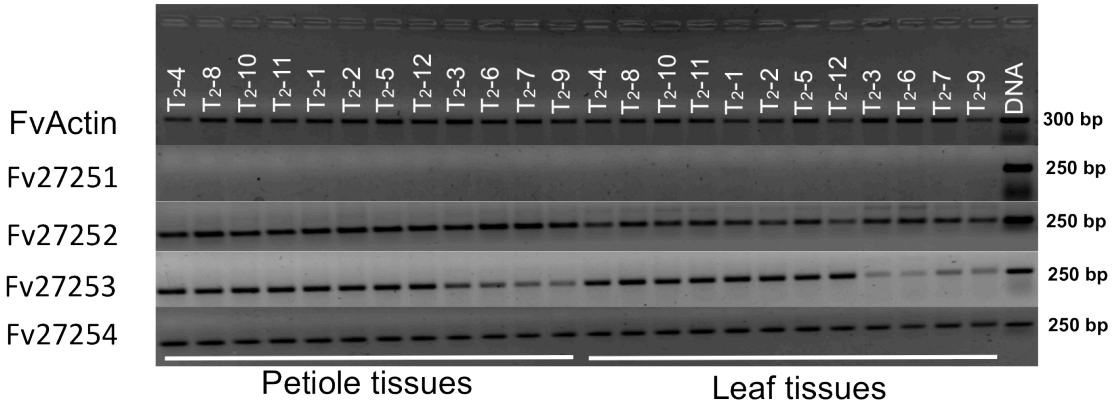
B)

Mutant Wild type

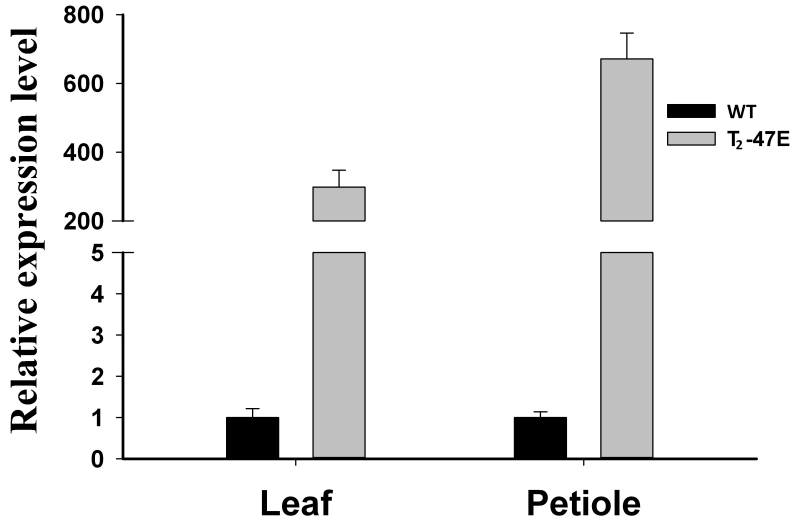
C)



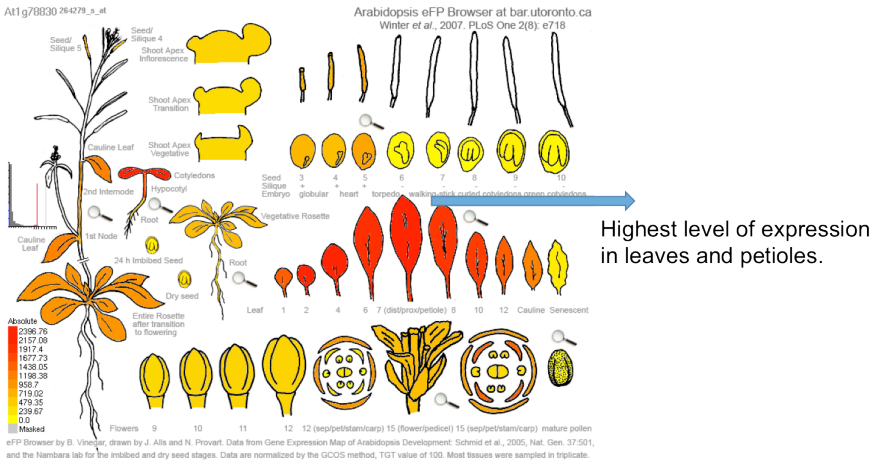
D)


T2-4
T2-8
T2-10
T2-11
T2-1
T2-2
T2-5
T2-12
T2-3
T2-6
T2-7
T2-9
T2-4
T2-8
T2-10
T2-11
T2-1
T2-2
T2-5
T2-12
T2-3
T2-6
T2-7
T2-9
DNA
FvActin
Fv27251
Fv27252
Fv27253
Fv27254
300 bp
250 bp
250 bp
250 bp
250 bp
Petiole tissues
Leaf tissues

E)


Relative expression level
800
600
400
200
5
4
3
2
1
0
WT
T2-47E
Leaf
Petiole

F)


Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
Highest level of expression
in leaves and petioles.

G)

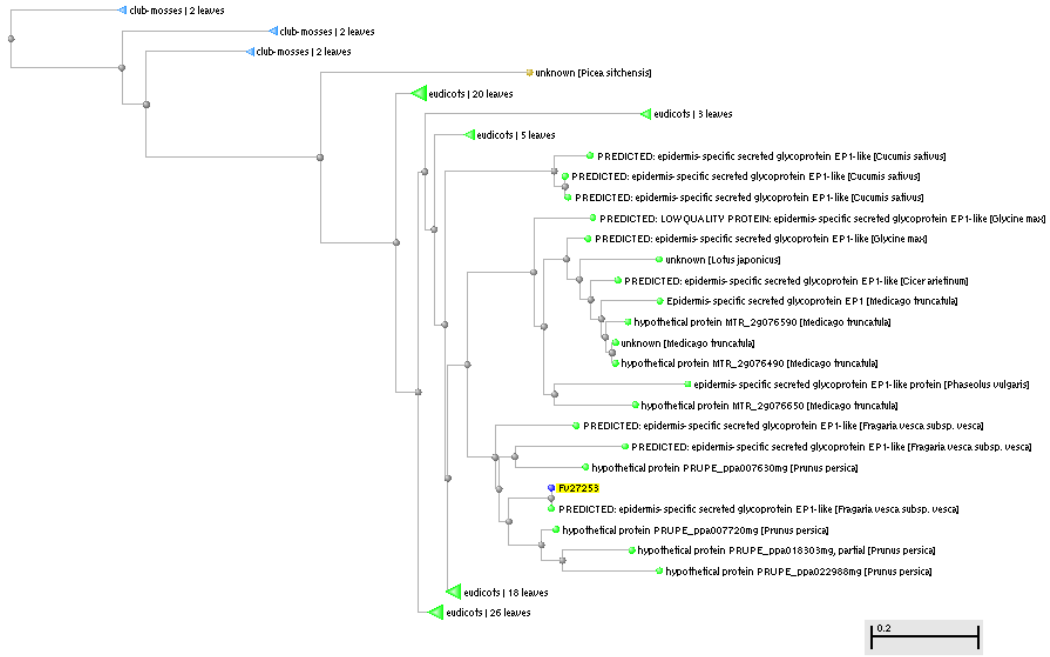


Figure 8 Analysis of an activation tagged strawberry mutant. **a** T_2 segregation analysis of 15 plants derived from T_1SH47E-7. Top band, PCR product between T-DNA specific and gene specific primers, indicating T-DNA insertion; lower band, PCR product of gene specific primers, indicating homozygote or hemizygote T-DNA insertion. T_1 transposant (lane 2), wild type strawberry (lane 3) and PCR master mix (lane 4) were positive and negative controls. In 15 plants tested (T_2-1 - 15, lanes 5-19), four homozygotes were found (T_2-4, 8, 10, 11) as well as six hemizygotes (T_2-1, 2, 5, 12, 14, 15) and five wild types (T_2-3, 6, 7, 9, 13). M, HyperLadder^{TM} II. **b** Phenotype of T_2 homozygote (left) and wild type strawberry (right). **c** Insertion site of *Ds* in the genome. Arrow shows direction of 4x35S enhancer **d** semi-quantitative RT-PCR test of four candidate genes upstream and downstream of *Ds* insertion. *FvActin* was used as control primer. Gel image was inverted for better quality. **e** Real time qRT-PCR to quantify gene expression level change of gene27253 between wild type and homozygote strawberry transposant. **f** eFP browser showing expression of gene27253 homolog in Arabidopsis, AT1G78860, in different tissue types. **g** Treeview of gene27253 and its homolog in other species. Method used here was neighbor joining and max sequence difference was set at 0.85.

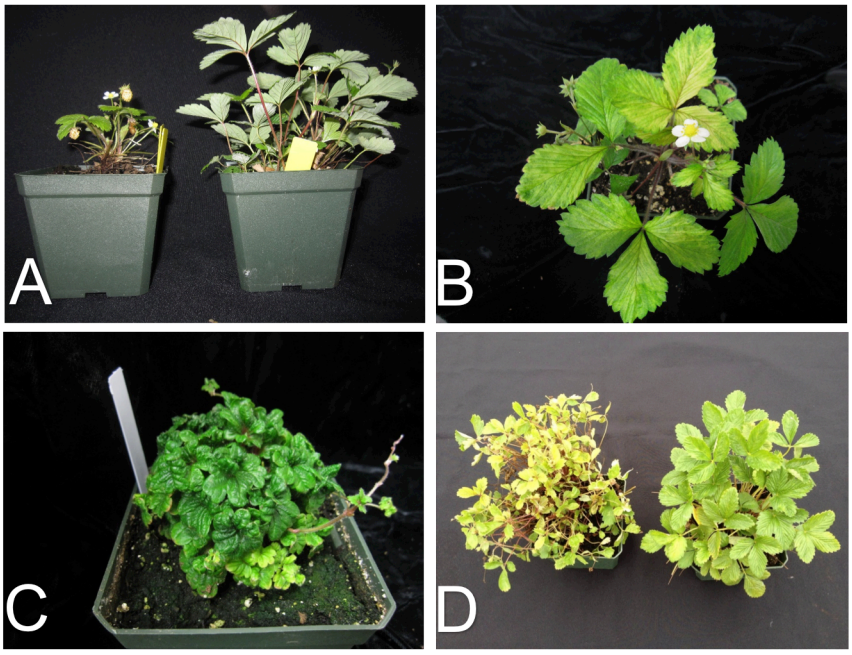


Figure 9 Phenotypic strawberry mutants collected during screening. **a** dwarf mutant (mutant on the left, wild type plant on the right). **b** golden leaf, normal flower and fruit set. **c** compact and high ploidy. **d** small leaves and small fruits mutant (mutant on the left, wild type plant on the right).

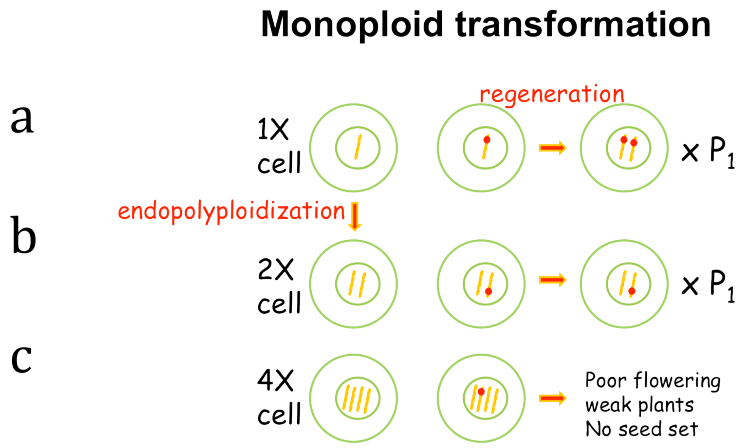
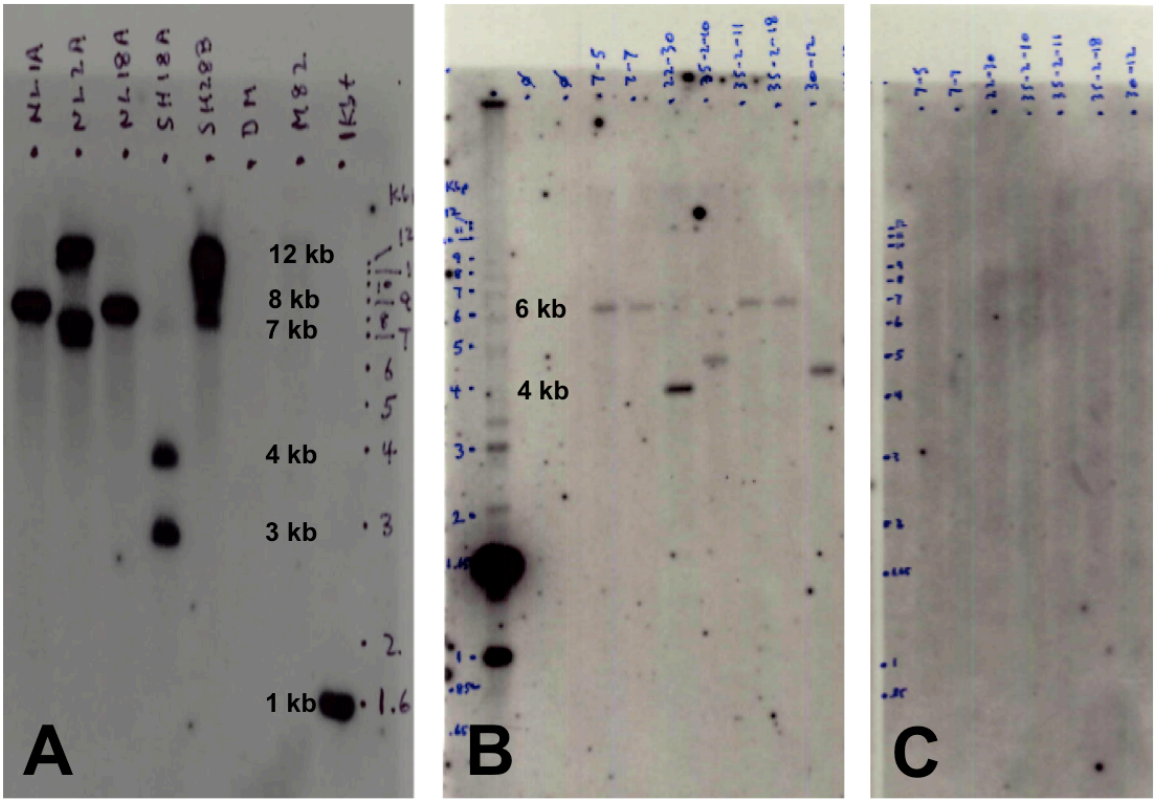


Figure 10 Transformation of monoploid potato and T-DNA insertion events. **a** T-DNA insertion occurred in monoploid cells and diploid regenerated plants would be homozygote for T-DNA. **b** T-DNA insertion happened in diploid cells after endopolyploidization, and then regenerated plants were hemizygote for T-DNA insertion. **c** T-DNA insertion occurred in tetraploid cells and no or few seed sets were found after crosses.

Supplemental materials



Supplemental Figure 1 Southern blot of *Ac-Ds*ATag-Bar_gosGFP primary strawberry regenerants and T_1 potato transposants. **a** Strawberry genomic DNA was extracted and digested with *HindIII* before hybridizing with probes designed from GFP gene. NL1A, NL18A had single insertion while NL2A and SH18A had two insertions, SH28B have multiple insertions as shown in picture. DM and M82 were negative wild type potato and tomato controls. **b** and **c** Seven potato transposants were hybridized using BAR (**b**) and GFP (**c**) probe. They were all single *Ds* insertions (**b**) without *Ac* (**c**). Four of them were same insertion and three were unique transposants as shown in (**b**).

Supplemental Table 1 Insertion sites of T_0 (A) and T_1 (B) strawberry mutants

A)

ID	TAIL primer	LG	Genomic region
SH63D	Ds5-3	chr03:24871009	Within gene24752
SH24A	Ds5-3	chr05:1271915	Within gene32452
SH57F	Ds5-3	chr06:25323921	Intergenic region
SH94B	Ds5-3	chr03:8185912	Within gene28868
NL-1A	Ds5-3	chr07:20817700	Within gene12548
NL-2A	LB-2a	chr03:24895143	Intergenic region
SH20A-2	LB-2a	chr01:1554801	Intergenic region
SH20	LB-2a	chr01:1452088	Intergenic region
SH23A-1	LB-2a	chr06:30306871	Promoter region of gene01132
SH26A	Ds5-3	chr01:3817994	Promoter region of gene30909
SH31B-1	LB-2a	chr03:29318831	Intergenic region
SH36N	LB-2a	chr04:11051429	Intergenic region
SH55A	Ds5-3	chr07:12522876	Intergenic region
SH58A	Ds5-3	chr04:17412685	Within gene04498
SH28K	Ds5-3	chr05:811328	Intergenic region
SH36C	Ds5-3	chr02:25759866	Within gene11679
SH16L	Ds5-3	chr02:8010707	Within gene17174
SH35A	Ds5-3	chr04:9083550	Intergenic region
SH73A	Ds5-3	chr01:740285	Within gene10040
SH47A	Ds5-3	chr03:11143356	Intergenic region
SH55A	Ds5-3	chr02:18032582	Within gene27738
SH88B	Ds5-3	chr03:24871025	Within gene27452
SH28	Ds5-3	chr02:35157079	Promoter region of gene15038
SH30	Ds5-3	chr07:22290344	Promoter region of gene12904
SH5B	Ds5-3	chr06:36098828	Intergenic region
SH17A	Ds5-3	chr03:5960616	5'UTR of gene20011
SH64C	LB-2a	chr04:21456703	Intergenic region
SH69D	LB-2a	chr07:7905548	Within gene04868
SH50B	LB-2a	chr02:4636738	Intergenic region
SH22A	LB-2a	chr02:9171115	Intergenic region
SH31A	LB-2a	chr03:29318943	Intergenic region
SH39A	LB-2a	chr07:16870084	Within gene26480
SH72A	LB-2a	chr04:7033563	Within gene29883
SH70A	LB-2a	chr03:28663161	Promoter region of gene20540
SH27B	LB-2a	chr06:35778706	3'UTR of gene09632
50B	LB-2a	chr02:4636761	Intergenic region
SH45B	LB-2a	chr01:8128274	Within gene13049

B)

ID	TAIL primer	LG	Genomic region
26G-6	Ds5-3	chr01:15648786	3'UTR of gene05608
T_1SH26G-6	Ds5-3	chr01:15648945	3' UTR of gene05608
T_1SH32A-14	Ds5-3	chr02:4124297	5'UTR of gene25444
T_1SH46C-1-2	Ds5-3	chr02:24381146	Within gene08583
67A	Ds5-3	chr03:606527	Within gene19469
T_1SH79A-2	Ds5-3	chr04:21607166	Within gene11991
T_1SH5D-1	Ds5-3	chr03:26261318	Within gene 05061
T_1SH5D-3	Ds5-3	chr03:26261286	Within gene 05061
T_1SH47E-7	Ds5-3	chr03:11143315	Promoter region of gene27253
T_1SH28B-6	Ds5-3	chr06:28261365	Within gene28794
$T_1$46C-2	Ds5-3	chr07:4715537	Within gene19244
T_1SH52B-1	Ds5-3	chr06:28261431	Within gene28794
G1-12	Ds5-3	chr03:11143356	Promoter region of gene27253
G1-103	Ds5-3	chr06:30307051	Promoter region of gene01132
T_1SH23M-2	Ds3-3	chr06:7290858	Promoter region of gene13817
T_1SH29D-13	Ds5-3	chr06:7682025	Promoter region of gene13488
T_1SH46C-1	Ds5-3	chr06:38632435	Within gene28400
T_1NL1A-P1A6	Ds5-3	chr05:15945225	Within gene27098
T_1NL1-P2E11	Ds5-3	chr03:9778113	Within gene14394
T_1SH23C-15	Ds3-3	chr01:11665100	Within gene24101
T_1SH26G	Ds3-3	chr07:17641516	Intergenic region
T_1SH29D-18	Ds5-3	chr07:18614019	Within gene20931
NL14	Ds5-3	chr07:20529794	Promoter region of gene21230
SH14C	Ds5-3	chr07:20530000	Promoter region of gene21230
NL15	Ds5-3	chr07:20529989	Promoter region of gene21230
$T_1$88B-3	Ds5-3	chr05:13591385	Within gene12161
T_1SH47E	Ds5-3	chr05:13591450	Within gene12161
T_1SH32A-13	Ds5-3	chr06:29366732	Within gene25930
SH23C-18	Ds5-3	chr01:6507634	Within gene12645
T_1SH52B-3	Ds5-3	chr04:19103467	Promoter region of gene04047

Supplemental Table 2 Insertion sites of T_0 (A) and T_1 (B) potato mutants, primers used for HiTAIL-PCR and genomic loci of each T-DNA in genome browser

A)

Plant	TAIL primer	LG	Genomic region
T_0-11B	Ds5-3	chr07:11737414	Intergenic region
T_0-36B	LB-2a	chr11:39899210	Within gene model PGSC0003DMT400047772

B)

Plant	TAIL primer	LG	Genomic region
T_1 58-77	Ds5-3	chr11:15045547	Intergenic region
T_1 58-89	Ds5-3	chr11:39929619	Within gene model PGSC0003DMT400047787
T_1 56-51	Ds5-3	chr11:35385869	Within gene model PGSC0003DMT400019014
T_1 57-87	Ds5-3	chr11:39283714	Intergenic region
T_1 56-35	Ds5-3	chr11:39949916	3'UTR of gene model PGSC0003DMT400048199
T_1 57-7	Ds5-3	chr11:39977543	Within gene model PGSC0003DMT400048108
T_1 57-10	Ds5-3	chr11:40728455	Intergenic region
T_1 57-1	Ds5-3	chr11:39978504	Within gene model PGSC0003DMT400048108
T_1 56-83	Ds5-3	chr11:40788523	Within gene model PGSC0003DMT400070475
T_1 58-1	Ds5-3	chr11:39979894	Within gene model PGSC0003DMT400048108
T1 57-11	Ds5-3	chr11:40689207	Intergenic region
T_1 58-103	Ds5-3	chr11:39988104	Promoter of gene model PGSC0003DMT400048202
T_1 56-20	Ds5-3	chr11:7704912	Intergenic region
T_1 58-116	Ds5-3	chr11:40033953	Within gene model PGSC0003DMT400048117
T_1 56-46	Ds5-3	chr11:40040769	3'UTR of gene model PGSC0003DMT400048181
T_1 56-1	Ds5-3	chr08:25800323	Within gene model PGSC0003DMT400005625
T_1 58-18	Ds5-3	chr11:41407038	Within gene model PGSC0003DMT400077713
T_1 58-104	Ds5-3	chr02:12722691	Within gene model PGSC0003DMT400059612
T_1 58-98	Ds5-3	chr11:39506716	Intergenic region
T_1 57-88	Ds5-3	chr11:39899250	Within gene model PGSC0003DMT400047772
T_1 56-25	Ds5-3	chr11:39978850	Within gene model PGSC0003DMT400048108
T_1 58-44	Ds5-3	chr11:35385284	Within gene model PGSC0003DMT400018991
T_1 58-115	Ds5-3	chr11:40656758	Promoter of gene model PGSC0003DMT400007810
T_1 57-45	Ds5-3	chr11:39929619	Within gene model of PGSC0003DMT400047787
T_1 56-18	Ds5-3	chr11:40777182	Intergenic region
T_1 56-76	Ds5-3	chr11:39937706	Within gene model PGSC0003DMT400047788
7-7	Ds5-3	chr07:2024762	5'UTR of gene model PGSC0003DMT400008300
22-30	Ds5-3	chr01:72210778	Intergenic region
T_1-14A	Ds5-3	chr02:16044607	Intergenic region

Supplemental Table 3 Primers sequences used in this study and their function.

Primer ID	Primer sequences (5'-3')	Purpose
G38BarF	CTGAAGTCCAGCTGCCAGAAACC	Multiplex PCR
G38BarR	CTGCACCATCGTCAACCACTACAT	Multiplex PCR
G38HPT5F	TCACAGTTTGCCAGTGATAC	Multiplex PCR
G38HPT5R	ATCGTTATGTTTATCGGCAC	Multiplex PCR
FvAAT3F	GTGACTTGGTTAACTTGCTC	Multiplex PCR
FvAAT3R	AAATTAGTCCAGCTCGTGAA	Multiplex PCR
SGT1-3F	TCCCTTGGACAGTAGATATTGCTG	Multiplex PCR
SGT1-3R	TTCCCAATCCCCTAACCTCG	Multiplex PCR
G38EDSF	GGAAACGACAATTTGATCTTTAGG	EDS PCR
G38EDSR	CTCAGTGGTTATGGATGGGAGTTG	EDS PCR
LAD1-1	ACGATGGACTCCAGAGCGGCCGCVNVNNNGGAA	HiTAIL-PCR
LAD1-2	ACGATGGACTCCAGAGCGGCCGCBNBNNNGGTT	HiTAIL-PCR
LAD1-3	ACGATGGACTCCAGAGCGGCCGCVVNVNNNCCAA	HiTAIL-PCR
LAD1-4	ACGATGGACTCCAGAGCGGCCGCBDNBNNNCGGT	HiTAIL-PCR
AC1	ACGATGGACTCCAGAG	HiTAIL-PCR
G38LB0a	CTTGATTTGGGTGATGGTTCACGTAGTGGG	HiTAIL-PCR
G38LB1a	GCAATCAGCTGTTGCCCGTCTCACTGGTG	HiTAIL-PCR
G38LB2a	ACCACCCCaGTAcATTAAAAACGTCCGCAATG	HiTAIL-PCR
Ds3-1	ACCCGACCGGATCGTATCGGT	HiTAIL-PCR
Ds3-2	ACGATGGACTCCAGTCCGGCCCGATTACCGTATTTATCCCGTTC	HiTAIL-PCR
Ds3-3	GTATTTATCCCGTTCGTTTTCGT	HiTAIL-PCR
Ds5-1	ACGGTCGGGAAACTAGCTCTAC	HiTAIL-PCR
Ds5-2	ACGATGGACTCCAGTCCGGCCCCGTTTTGTATATCCCGTTTCCGT	HiTAIL-PCR
Ds5-3	TACCTCGGGTTCGAAATCGAT	HiTAIL-PCR
T2-5DF	GTCTATTTAGTACTAATTTAAACCCATTGTTGCT	Zygosity test
T2-5DR	CGCAGTCGCCGTCAACGA	Zygosity test
T2-47EF	GCCCCGTGTTTGAATATCACAT	Zygosity test
T2-47ER	GGGACTTGAGCCTGAGCAAGA	Zygosity test
T2-23CF	TGAGTGGGTTAGGTTTGCCCG	Zygosity test
T2-23CR	GTCCAACCGTCATAATCTGGTACATG	Zygosity test
FvActinF	GGTCTCGAACATTATCTGGGTCAT	RT-PCR
FvActinR	AGGCCGGGTTTGCTGGAG	RT-PCR
Fv27251F	CCAATGGCAACATGGTACTCTACG	RT-PCR
Fv27251R	GCCGATGACTGTAATGTTAAGTCCC	RT-PCR
Fv27252F	CGACGCCCCGGTGAGGAG	RT-PCR
Fv27252R	GGCGCTACATGGAGTCGTTGAGT	RT-PCR
Fv27253F	CCCCGAATGCCTTCACTTTGG	RT-PCR
Fv27253R	CGCCCTTGGAGTTGTAGAGGACC	RT-PCR
Fv27254F	CAAGGCGCGTTATGGTGCATG	RT-PCR
Fv27254R	GGCTCACAGCTCTCACTCCAACCC	RT-PCR
Fv27255F	GGATGCAAACCGAAACCACCC	RT-PCR
Fv27255R	ATGGGGGTAGTCAAAGCTTTGCC	RT-PCR

3. NEW STRATEGIES FOR TRANSPOSON TAGGING BY USING GAMETOPHYTE SPECIFIC PROMOTERS

Abstract:

Transposon tagging has been an important approach in studying gene function in plant genomics. Several strategies have been applied but the obstacle of somatic transposition has been a great challenge. In order to attempt to reduce somatic transposition and increase germinal transposition in crop development, we synthesized two optimized transposon-tagging constructs, where the transposase gene was controlled by one of two microspore-specific promoters and the *Ds* element harbored both kanamycin resistance as well as promoters to control transposase and green fluorescent protein. Positive selection for kanamycin resistance and negative selection for GFP fluorescence would be expected to yield stable transposon tagged lines. Both constructs were transformed into diploid potato P1 clone via *Agrobacterium*-mediated transformation. Transformed potato plants were out-crossed to as pollen parents to wild type diploid potato plants to generate T_1 putative transposants. T_1 seeds were germinated in B5 basal medium with 50 mg/L kanamycin and screened for GFP before transferring to soil. Anther culture was also performed on transgenic potatoes to observe possible transposition during the androgenetic process. Multiplex PCR was used to verify the selection strategy for *Ds* only or EDS+*Ds* plants and HiTAIL-PCR used to amplify flanking sequences adjacent to *Ds* insertion sites. Transposants were found from anther culture derived plants but not from T_1 plants, indicating promoter specificity change in heterologous genome. Our results provide resource for further study on improving transposon-tagging efficiency.

Introduction

Transposon tagging has been a powerful tool in studying plant genomics, as an alternative to traditional T-DNA insertional mutants. Unlike using T-DNA, many mutants can be theoretically generated from a few transgenic launch pads. Transposon-tagged mutant collections have been established in many plant species, such as Arabidopsis (Ito et al. 2005; Kuromori et al. 2004), maize (Brutnell and Conrad 2003), rice (Greco et al. 2003), barley (Ayliffe et al. 2007), poplar (Fladung and Polak 2012), maize (Vollbrecht et al. 2010), soybean (Mathieu et al. 2009), etc. The success of generating large transposant populations in plants is usually dependent on

efficient transformation as well as easy and economical screening for transposants using different marker genes, such as GFP, herbicide resistance, or antibiotic resistance, etc. Transposition efficiency varies among different species. Previous studies have been focused on controlling transposase gene expression to reduce somatic transposition, such as using meiotic specific promoter from lily to control transposase expression resulting in only a 1% increase in the transposition frequency in rice (Morita et al. 2003), or using heat shock promoter to induce transposase expression at generative growth stage of plants (Nishal et al. 2005). These results have resulted in some improvement compared with constructs using constitutive promoters, such as p35S or ubiquitin promoters. The timing and tissue specificity for transposase gene expression might be important in improving transposition efficiency. Controlling transposase gene expression using tissue specific promoters is less labor intensive than using inducible promoters, especially for species where flowering is not synchronized. So in order to improve germinal transposition, one of the best strategies would be to repress transposase expression during the vegetative growth stage and only initiate expression from generative stage, such as flowering. Choosing promoters active in male gametogenesis is more promising than using gynogenesis specific promoters, because the number of pollen grains generated from each flower are a lot more than eggs generated in dicots like potato and strawberry, thus more chance of getting unique transposants.

During pollen generation and maturation in higher plants, four microspores are generally formed from each pollen mother cell after two meiosis stages, meiosis I and meiosis II. Chromosome rearrangement may occur during these stages. Microspores are released and bicellular pollen grains formed from each microspore after mitosis. Each pollen grain contains one vegetative cell nucleus and one generative cell. The generative cell then undergoes mitosis to produce two sperm cells. The vegetative cell becomes a companion cell that helps sperm cells to finish double fertilization (Alonso et al. 2003; Khurana et al. 2013). If transposition occurs after microspores are formed, then the sperm cells coming from each microspore might have unique *Ds* reinsertions, even though half of seeds germinated should be wild type. However, all of the stages of pollen maturation are brief so the timing of transposase expression is critical.

Due to the limited availability of candidate microspore-specific promoters characterized in potato, we tried to find alternatives in other related dicot species. Many plant tissue specific promoters have been identified, some of which have specific timing for expression. In the study

in Arabidopsis, the sidecar pollen gene, which encoded a male gametophyte-specific protein that regulated correct patterning and asymmetric divisions, was found to be expressed at a specific timing during pollen development (Clough and Bent 1998; Oh et al. 2010; Ramachandran and Sundaresan 2001). A knockout mutant of the sidecar pollen gene was found to have a four-cell microspore structure instead of three. In a study in *Antirrhinum*, the DEFH125 MADS-box protein that was found to be participating in pollen development, was further characterized as a transcription factor that interacted with other MADS-box genes, DEF/GLO; a Gus (β-glucuronidase) staining experiment confirmed the onset of promoter activity starting at late unicellular microspore stage (Lauri, Xing et al. 2006). Promoters from both of these genes have specific domains that are important in determining their function, such as CArG-boxes.

Another factor in transposon tagging in diploid crop species with low transposition efficiency is that the particular transposase gene was originally cloned from the monocot species maize. It has been reported that genes function differently in heterologous genomes between monocot and dicot species, probably due to promoter specificity (Assem et al. 2001; Khurana et al. 2013). In order to overcome this, we sought to enhance expression of the transposase gene by switching to dicot preferred codon sequences based on online codon usage databases without changing amino acid composition.

The transposase gene has five exons and four introns, which compose two reading frames. Reading frame A contains five exons and reading frame B contains four exons (exons 2 to 5). Both reading frames were active for gene expression. The *pAcDs* construct used in strawberry transposon tagging population had reading frame b of transposase (Veilleux et al. 2012), whereas the *Ac-Ds*ATag-Bar_gosGFP construct used in strawberry, potato (chapter 2) had reading frame a, both of which were functional in these diploid species but each was dogged by the problem of somatic transposition.

Here we report the study of using the two new *AcDs* constructs where transposase gene expression was controlled by either of two candidate promoters, sidecar pollen from Arabidopsis (AtSCP) and AmDEFH125 from *Antirrhinum* (AmDEFH125), in order to attempt to enhance germinal transposants in potato. We used both anther culture and cross-hybridizations as potential test systems for transposition.

Methods and materials

Vector construction

Transposase (Genbank Accession X05424), GFP (Genbank Accession HQ702996), and *NPTII* (Genbank Accession AY237649.1) gene sequences were optimized for diploid crop species based on codon usage database (http://www.kazusa.or.jp/codon/). Optimized sequences were shown in supplemental figure 2. Three fragments were synthesized at Genscript, Inc: 1) tpase-pUC57 fragment containing transposase gene; 2) GFP-pUC57 fragment containing GFP gene; 3) *Ds*-pUC57 fragment containing inverted repeat (IR), pmasPro-*NPTII*-nosT as well as LB-0a and LB-1a TAIL primer sequences adjacent to IR that could facilitate HiTAIL-PCR. All of the three genes were cloned in pUC57 cloning vector.

To synthesize the whole cassette, we first cut the transposase gene from the tpase fragment by performing *AvrII-XhoI* double digestion and ligating transposase into the GFP-pUC57 fragment, forming a tpase-GFP-pUC57 fragment. Arabidopsis Sidecar pollen promoter (Oh et al. 2010) and snapdragon DEFH125 promoter (Lauri et al. 2006) were amplified from wild type *Arabidopsis thaliana* ecotype Col-0 and *Antirrhinum majus* "Rocket Mix" seedlings by PCR using primers and conditions mentioned in respective publications. Amplified promoters were inserted in the *Ds*-pUC57 fragment using *PstI-NotI* double digestion and ligation, forming Promoter-*Ds*-pUC57 fragment. The Promoter-*Ds* part was cut from pUC57 vector by *XhoI-NgoMIV* double digestion and then inserted in tpase-GFP-pUC57 fragment, forming the tpase-promoter-*Ds*-GFP-pUC57 fragment. Finally the whole T-DNA cassette (tpase-Promoter-*Ds*-GFP) was digested by *SalI* and inserted into pCAMBIA0380 vector. The two vectors were transformed into *Agrobacterium* strain GV2260 for plant transformation. Two constructs were synthesized this way, *pAcDs-AtSCP* and *pAcDs-AmDEFH125* (Fig. 11).

Plant material and *Agrobacterium*-mediated transformation

Potato transformation was modified from An et al. (1986). *In vitro* plantlets of diploid *Solanum tuberosum* Group Phureja clone P1, which is closely related to the sequenced doubled monoploid (DM), were established from existing plants (Pyrex® 25 mm × 150 mm test tubes containing 20 mL of MS basal medium) for 2 weeks before leaf tissue was excised and wounded by slicing using a scalpel blade prior to *Agrobacterium*-mediated transformation. Sliced leaves were placed

adaxial-side down in petri dishes containing 40 mL callus induction medium (MS salts, 0.9 mg/L thiamine HCL, 3% sucrose, 0.8 mg/L zeatin riboside, 2 mg/L 2,4-D, pH5.8, 2g/L Phytagel) for 2 days at 22°C. *Agrobacterium* was grown for 16 h and centrifuged at 4000 rpm, 4°C for 20 min to collect pellets. Bacteria pellets were suspended in 20 mL dilution medium (MS salts with vitamins, 3% sucrose, pH5.5) and leaf explants were immersed in dilution medium for 20 min before moving back to callus induction medium for 5 days. Leaf explants were washed five times with wash-off medium (MS salts with vitamins, 0.9 mg/L thiamine HCl, 3% sucrose, 250 mg/L cefotaxime, pH5.8) and moved to shoot induction medium (MS salts with vitamins, 0.9 mg/L thiamine HCl, 3% sucrose, 0.8 mg/L zeatin riboside, 4 mg/L GA3, 250 mg/L cefotaxime, 25 mg/L kanamycin, pH 5.8, 2 g/L phytagel). Regenerated plants were screened for GFP expression and DNA was extracted for multiplex PCR to confirm T-DNA presence.

Flow cytometry

Transgenic potato plants were processed for flow cytometry to estimate ploidy. Sample preparation was according to Owen et al. (1988). Approximately 0.5 g leaf tissue from *in vitro* plants was placed in a glass petri dish and 1 mL of chopping buffer (3.52 mg/ml sodium citrate, 1.7 mg/ml MOPS, 3.7 mg/ml $MgCl_2$, 0.04% v/v Triton X-100) was added; 0.5 mL of filtrate from chopped tissue was pipetted into a microcentrifuge tube; 0.25 mL of ribonuclease (0.8 mg/mL) was then added to each tube and incubated at room temperature for 30 min. PI stain (0.125 mL of 0.4 mg/mL propidium iodide) was added to each sample and incubated on ice for 30 min before running flow cytometry. Samples were run in FACSCalibur flow cytometer (BD Biosciences) using Cellquest Pro acquisition software (BD Biosciences). Results were analyzed using FlowJo software (Treestar, Inc).

Anther culture

Anther culture process was according to Snider and Veilleux (1994). Flower buds containing microspores at the late uninucleate stage (2.5-4 mm long anther for diploid potato flower, 3-5 mm long anther for tetraploid potato flower) were collected, placed in moist paper towels at 4°C in the dark for 3 days. Buds were then sterilized by immersing them for 1 min in 80% ethanol, then 5 min in full strength bleach with Tween 20, and finally rinsed three times in sterile distilled water. Flower buds were then dissected and anthers were placed in 125 ml Delong culture flasks

with 15 ml liquid culture media (half-strength MS basal medium, 0.4 mg/L thiamine HCl, 60 g/L sucrose, 2.5 g/L activated charcoal, 2.5 mg/L BA, and 0.1 mg/L IAA, pH 5.8). Flasks were placed on a shaker (100 rpm) at room temperature. After 5 weeks, embryos were harvested and transferred to regeneration medium (3.2 g/L Gamborg's B5 basal salt, 50 mg/L $CaHPO_4$, 748 mg/L $CaCl_2$, 250 mg/L NH_4NO_3, 10 g/L sucrose, 2 g/L phytagel, 0.5 mg/L GA_3, pH 5.6). Embryos were incubated at 22°C/18°C under 16 h photoperiod. Embryos that converted into plantlets were transferred to tubes containing 20 ml MS basal medium. The regenerated plants were transferred to a growth chamber at 22°C/18°C under a 16 h photoperiod with light intensity of 175 μm/m^2/s.

Out-crosses for fruit set

T_0 transgenic potato as well as wild type *S. tuberosum* Group Tuberosum RH89-039-16 clone (RH) and DMRH (F_1 hybrids between DM and RH) plants were grown in a greenhouse under long day conditions in two consecutive seasons (fall 2012 and spring 2013) for pollen collection and crosses. Pollen collected from diploid transgenic potato line AtSCP-5C-1 containing *pAcDs-AtSCP* was crossed with wild type diploid RH or DMRH plants to generate fruits in order to screen for candidate T_1 transposants.

Seed germination

Harvested T_1 seeds were surface sterilized by first washing with 80% ethanol for 5 sec, and then washing with 30% bleach with Tween 20 for 20 min. Seeds were then rinsed twice with distilled water and soaked in 500 uL of 1,000 ppm GA_3 overnight. Next day seeds were washed again with distilled water three times and transferred into 125 mL flasks containing 15 mL of B5 basal medium (B5 basal salt, 30 g/L sucrose, pH 5.8) and 50 mg/L kanamycin. Flasks were incubated on a shaker at 80 rpm under light for 2-4 weeks before germinated seeds were selected for kanamycin resistance and transferred to soil.

Plant growth conditions

T_1 plants were maintained in a growth chamber (Conviron, model ATC60) under long day conditions (16 h/8 h day/night, 22 °C/18°C, light intensity 175 μm/m^2/s) for 3-4 weeks and leaf tissue from plants was collected for DNA extraction.

Multiplex PCR

Multiplex PCR was performed to screen for putative transposants. Multiplex PCR was carried out using 2x Immomix red polymerase mix (BIO-25022) plus 1 mM $MgCl_2$, 0.2 mM of each primer and 50 ng of DNA in 25uL final volume. Primers in this multiplex PCR were designed to amplify the empty donor site (EDS: tpaseR1 and EGFPR1), the full donor site (FDS: p35SF1 and EGFPR1), *NPTII* (NPTIIF2 and NPTIIR2), *SQE* (squalene epoxidase, SQEF and SQER) (Supplemental Table 4). The PCR program was one cycle (95 °C for 10 min), 35 cycles (94 °C for 30 s, 56 °C for 30 s, 72 °C for 1.5 min), and a final extension cycle of 5 min at 72 °C.

HiTAIL-PCR

To determine insertion sites, we performed HiTAIL-PCR on both launch pads and candidate transposants following the same protocol (Liu and Chen 2007). For T_0 plants, HiTAIL-PCR was performed using primers from the right border of the *Ac* element of the T-DNA (RB-TAIL). For candidate transposants, HiTAIL-PCR was performed using primers from the inverted repeat of the *Ds* element, *Ds*5'IR (*Ds*5 TAIL). R1 products were loaded in 1% agarose gel and selected bands were cut, purified and sequenced before BLAST searching into the potato genome browser (http://solanaceae.plantbiology.msu.edu/integrated_searches.shtml) to determine insertion loci.

Results

Transformation efficiency

For each construct, five independent transformation experiments were tried out and in each experiment five leaves were used for inoculation. Each leaf was then cut into four secondary explants for regeneration. Six lines of transgenic potatoes with the *pAcDs-AtSCP* construct and 13 lines with the *pAcDs-AmDEFH125* construct were verified after screening approximately 200 regenerated and putatively transformed plants using multiplex PCR. Among the six transgenic lines carrying *pAcDs-AtSCP,* two lines, AtSCP-5D-1 and AtSCP-B11, showed early somatic transposition with EDS products. Among the transgenic lines carrying *pAcDs-AmDEFH125*, one line, AmDEFH125-199 showed early somatic transposition with EDS, FDS and *NPTII*, which also suggested that it had more than one T-DNA insertion in the genome (Fig. 12, Table 5).

Ploidy of T_0 transgenic potatoes

Flow cytometry was used on primary regenerated transgenic potato to estimate ploidy (Fig. 13). In plants carrying *pAcDs-AtSCP*, four were diploid and two were tetraploid; for plants carrying *pAcDs-AmDEFH125*, six were diploid and seven were tetraploid.

Anther culture derived plants

Anther-derived plants (n=123) were obtained from two T_0 transgenic lines, AtSCP-5C-1 and AmDEFH125-104, and screened for possible transposition (Fig. 14). Multiplex PCR revealed a total of 32 candidate transposants (EDS+*Ds* or *Ds* only, Table 6), from both lines. Plants derived from AmDEFH125-104 had more *Ds* only or EDS+*Ds* plants (61% against 5%), whereas the AtSCP-5C-1 had more wild type regenerated plants (86% vs. 26%) (Table 6). Because the anther-derived plants were most likely from cells at mid-or late-unicellular stage (Datta 2005), expression of transposase of both constructs might have started during this stage.

Ac/Ds footprint

In the 12 plants (both T_0 plants and anther culture derived plants) that showed EDS during multiplex PCR, the EDS PCR products were purified and sequenced to observe footprints. In the eight successful sequencing events, different patterns were found at the excision point between transposase and GFP (Fig. 15). The difference also indicated excision events occurred in different cells during either transformation or anther culture process.

T_1 screening by crosses

All transgenic potatoes were planted for two constitutive seasons (fall 2012, spring 2013) in a greenhouse under long day condition, however, only AtSCP-5C-1 and AmDEFH125-104 had normal flowers with enough pollen shed for crosses. Other plants either had abnormal flowers that withered prematurely or had normal looking flowers but not enough pollen could be collected. It could be because of transposition occurred in pollen development that caused abortion of pollens. AtSCP-5C-1 was crossed with diploid wild type RH or DMRH plants to generate T_1 seeds. A total of 25 fruits was collected. About 2500 seeds were germinated in liquid B5 basal medium containing 50 mg/L kanamycin. From theses, 550 kanamycin resistant and GFP negative plants (Fig. 16) were planted in soil 3 weeks after germination. In PCR screening

of 257 plants (59 DMRH × AtSCP-5C-1, 198 RH × AtSCP-5C-1), 241 were *AcDs* plants (51 DMRH × AtSCP-5C-1, 190 RH × AtSCP-5C-1) and 16 were wild type plants (8 DMRH × AtSCP-5C-1, 8 RH × AtSCP-5C-1); no putative transposants were found. The kanamycin selection was effective in eliminating wild type seedlings (only 6% were wild type escapes). GFP expression was insufficient to confidently separate GFP positive plants from GFP negative plants (Fig. 16B).

HiTAIL-PCR for T-DNA or *Ds* insertion site

A modified HiTAIL-PCR protocol was used to amplify flanking sequences from the right border of *Ac* (T_0 launch pads) or the inverted repeat region of *Ds* (transposants) to locate the insertion sites in the genome (Liu and Chen 2007). Eleven of the 19 T_0 insertion sites were anchored on different potato chromosomes by using RB primer (Table 7).

In amplification of flanking sequences from inverted repeat region of the 32 candidate transposants identified from anther culture derived plants as well as the two T_0 plants with EDS/*Ds*, we were able to locate two unique insertion sites from three transposants, as well as one insertion in the backbone of *pCAMBIA0380* vector (Fig. 17). HiTAIL-PCR results from other candidate transposants surprisingly amplified sequences from the inverted repeat region into the transposase gene, indicating an intact FDS. This could be due to a second T-DNA insertion site in the genome that we were not able to detect in multiplex PCR.

It came as a surprise that transposants could be found in anther culture derived plants but not from T_1 plants coming from seeds, out of the same T_0 line AtSCP-5C-1. We were also able to find a few mutants with interesting phenotype, such as curly leaf, which was from crosses between RH and AtSCP-5C-1 (Fig. 18).

Discussion

Forward genetics and transposon tagging

Using forward genetics method in characterizing gene function has a long and successful history for plant biology. The success of using forward genetics is mostly dependent on easy transformation and easy screening. The potato draft genome annotated over 30,000 candidate genes (The Potato Genome Sequencing Consortium 2011). In order to create single T-DNA insertional mutants to knockout each gene in potato, it would require great effort. An alternative

method is using the *AcDs* strategy. The advantage of using *AcDs* strategy is that only a few transgenic events were required to derive sufficient launch pads that can be used generate unlimited number of transposon tagged lines if independent germinal transposition can be induced. In most cases, a successful launch pad on each chromosome could significantly increase the chance of saturating the genome, because of the better opportunity of localized transposition near launch pads.

Marker gene selection efficiency

T_1 seeds from crosses between wild type RH or DMRH and AtSCP-5C-1 were germinated in liquid medium containing 50 mg/L kanamycin. Kanamycin selection efficiency was verified by multiplex PCR, and the result showed about 94% of the green seedlings contained the *NPTII* gene, only 6% of seedlings were wild type escapes. GFP selection was not as effective as kanamycin selection. GFP expression in elongating roots among transgenic and wild type potatoes was hard to identify.

In chapter 2, the GFP expression in *Ac-Ds*ATag-Bar_gosGFP construct was working for strawberry but not for potato. However, the *Ac* element has two selectable markers, hygromycin resistance and GFP, so future optimization of our present constructs should consider the possibility of double markers on the *Ac* element in order to increase selection efficiency, yet they do not necessarily have to be fluorescent markers.

Germinal transposition and tissue-specific promoters

In our lab we have generated two populations of transposon tagged mutants in different crop species, strawberry and tomato. A big challenge in screening for unique transposants has been the considerable number of somatic transpositions, meaning that transposition occurred prematurely, resulting in many different seedlings harboring the same transposition site. In order to avoid this, several efforts have been tried, such as using a two-component system or using tissue-specific or inducible promoters to control transposase expression. In this study, we proposed to use tissue specific promoters to drive transposase expression so that transposition would not occur until a certain stage during plant generative growth, such as pollen development. Pollen development and maturation in higher plants have seven major steps including two meiosis stages and one mitosis stage. If transposition occurred after meiosis, which meant it

occurred in each individual free microspore, then the chance of getting unique or global transposition would be better. For any single hemizygous T-DNA insertion, then half of the pollen grains would be wild type and the other half could carry unique *Ds* reinsertion. In addition, we also attempted to optimize nucleotide sequences of the transposase gene to be more suitable for diploid species based on codon usage database. In our literature study for microspore specific promoters in dicot species, we focused on two candidates, the sidecar pollen promoter from Arabidopsis and the promoter for MADS-box transcription factor DEFH125 from snapdragon. Both promoters were well characterized and functioned during the optimum developmental stage. We then did BLAST search using these two candidate genes in the potato genome browser to look for orthologous genes in potato, but it did not return good hits with high e-values, so we decided to use promoters from these two candidate genes instead. In the 19 lines we generated from *Agrobacterium*-mediated transformation, three lines exhibited early transposition. This indicates that the modified transposase gene was functional. Early somatic transposition is usually due to tissue culture stress to cells during transformation and regeneration. Somehow the tissue culture stage must have led to transcription of transposase even with the tissue-specific gametophyte promoters. Surprisingly out of the 19 lines we only observed normal flowers with abundant pollen shed in two lines, AmDEFH125-104 and AtSCP-5C-1. All the other lines had either early flower abortion or poor pollen shed. AtSCP-5C-1, which had normal flowers and abundant pollen grains, did not generate any transposants in the T_1 progeny. It is possible that expression of one of the transgenes or transposition itself was incompatible with normal developmental processes. In a study that used heat shock promoters to induce transposase expression in Arabidopsis, it was found to be effective but transposition seemed to be localized (Nishal et al. 2005). In an earlier study where multiple constructs were synthesized using inducible promoters to control transposase gene expression in rice, it was found that early somatic transposition was still inevitable (Qu et al. 2009). The idea of using native microspore specific promoters to induce transposition is less labor intensive compared to the efforts of using inducible promoters.

Potato transposants derived from anther culture

The result that transposants were found among anther-derived plants but not from natural crosses with wild type plants was not expected. The EDS results of different footprints from different

plants indicated that excision occurred in different cells. And the result of different transposants found in different anther-derived plants confirmed this. Anther culture is a technique used in many plant species for developing homozygous lines and has been used extensively in potato (Veilleux 1999) to develop monoploids and doubled monoploids from diploid germplasm or dihaploids from tetraploid germplasm. What differentiates the cell division process of microspores that undergo androgenesis from those that form pollen grains has usually been associated with an irregular equal rather than the normal unequal division (Iyer and Raina 1972). In many cases including ours, the pretreatments applied to flower buds destined for anther culture, such as temperature shock or nutrient starvation, are directed towards halting the postmeiotic mitosis in order to encourage the abnormal development. It may be that an extended duration of postmeitoic mitosis provides a greater opportunity for transposition to occur compared to the relatively quick passage of developmental stages during *in planta* microsporogenesis. One possible experiment to test this theory would be applying a cold shock to inflorescences prior to pollen collection to provide a greater opportunity for transposase activity. Previously another natural transposable element *mPing* has been observed to be active in anther culture of rice, *Oryza sativa* (Kikuchi et al. 2003), or even in cell culture (Jiang et al. 2003). This evidence supported our finding of transposase being active during anther culture.

Even though we did not find transposants from T_1 seeds, we did find transposants from anther culture derived plants for both lines we tested, and the transposition frequencies were significantly different. Further studies to better understand the mechanism causing the differences might be needed in the future.

Acknowledgements

We thank Dr. Bingyu Zhao from Horticulture department for kind help in codon optimizing and gene cloning process, and Jeffrey Burr for help in maintaining greenhouse plants.

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301:653-657

An G, Watson BD, Chiang CC (1986) Transformation of Tobacco, Tomato, Potato, and *Arabidopsis thaliana* Using a Binary Ti Vector System. Plant Physiol 81:301-305

Assem S, El-Itriby H, Saad M, Madkour M (2001) Comparison of the efficiency of some novel maize promoters in monocot and dicot plants. Arab J Biotech 5:57-66

Ayliffe MA, Pallotta M, Langridge P, Pryor AJ (2007) A barley activation tagging system. Plant Mol Biol 64:329-347

Brutnell TP, Conrad LJ (2003) Transposon tagging using *Activator (Ac)* in maize. Methods Mol Biol 236:157-176

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16:735-743

Datta SK (2005) Androgenic haploids: Factors controlling development and its application in crop improvement. Curr Sci India 89:1870-1878

Fladung M, Polak O (2012) *Ac/Ds*-transposon activation tagging in poplar: a powerful tool for gene discovery. BMC Genomics 13:61

Greco R, Ouwerkerk PB, De Kam RJ, Sallaud C, Favalli C, Colombo L, Guiderdoni E, Meijer AH, Hoge Dagger JH, Pereira A (2003) Transpositional behaviour of an *Ac/Ds* system for reverse genetics in rice. Theor Appl Genet 108:10-24

Ito T, Motohashi R, Kuromori T, Noutoshi Y, Seki M, Kamiya A, Mizukado S, Sakurai T, Shinozaki K (2005) A resource of 5,814 *dissociation* transposon-tagged and sequence-indexed lines of *Arabidopsis* transposed from start loci on chromosome 5. Plant Cell Physiol 46:1149-1153

Iyer RD, Raina SK (1972) The early ontogeny of embryoids and callus from pollen and subsequent organogenesis in anther cultures of *Datura metel* and rice. Planta 104:146-156

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR (2003) An active DNA transposon family in rice. Nature 421:163-167

Khurana R, Kapoor S, Tyagi AK (2013) Spatial and temporal activity of upstream regulatory regions of rice anther-specific genes in transgenic rice and *Arabidopsis*. Transgenic Res 22:31-46

Kikuchi K, Terauchi K, Wada M, Hirano HY (2003) The plant MITE *mPing* is mobilized in anther culture. Nature 421:167-170

Kuromori T, Hirayama T, Kiyosue Y, Takabe H, Mizukado S, Sakurai T, Akiyama K, Kamiya A, Ito T, Shinozaki K (2004) A collection of 11 800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. Plant J 37:897-905

Lauri A, Xing S, Heidmann I, Saedler H, Zachgo S (2006) The pollen-specific *DEFH125* promoter from *Antirrhinum* is bound in vivo by the MADS-box proteins DEFICIENS and GLOBOSA. Planta 224:61-71

Liu YG, Chen Y (2007) High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. BioTechniques 43:649-650, 652, 654 passim

Mathieu M, Winters EK, Kong F, Wan J, Wang S, Eckert H, Luth D, Paz M, Donovan C, Zhang Z, Somers D, Wang K, Nguyen H, Shoemaker RC, Stacey G, Clemente T (2009) Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository. Planta 229:279-289

McCormick S (2004) Control of male gametophyte development. Plant Cell 16 Suppl:S142-153

Morita R, Hattori Y, Yokoi S, Takase H, Minami M, Hiratsuka K, Toriyama K (2003) Assessment of utility of meiosis-associated promoters of lily for induction of germinal *ds* transposition in transgenic rice. Plant Cell Physiol 44:637-642

Nishal B, Tantikanjana T, Sundaresan V (2005) An inducible targeted tagging system for localized saturation mutagenesis in Arabidopsis. Plant Physiol 137:3-12

Oh SA, Park KS, Twell D, Park SK (2010) The *SIDECAR POLLEN* gene encodes a microspore-specific LOB/AS2 domain protein required for the correct timing and orientation of asymmetric cell division. Plant J 64:839-850

Owen HR, Veilleux R, Haynes F, Haynes K (1988) Variability for critical photoperiod for tuberization and tuber yield among monoploid, anther-derived genotypes of *Solanum phureja*. J Amer Soc Hort Sci 113:755-759

Qu S, Jeon JS, Ouwerkerk PB, Bellizzi M, Leach J, Ronald P, Wang GL (2009) Construction and application of efficient *Ac-Ds* transposon tagging vectors in rice. J Integr Plant Biol 51:982-992

Ramachandran S, Sundaresan V (2001) Transposons as tools for functional genomics. Plant Physiol Bioch 39:243-252

Snider K, Veilleux R (1994) Factors affecting variability in anther culture and in conversion of androgenic embryos ofSolanum phureja. Plant Cell Tissue Organ Cult 36:345-354

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189-195

Veilleux RE (1999) Anther culture of potato and molecular analysis of anther-derived plants as laboratory exercises for plant breeding courses. HortTechnology 9:585-588

Veilleux RE, Mills KP, Baxter AJ, Upham KT, Ferguson TJ, Holt SH, Lu N, Ruiz-Rojas JJ, Pantazis CJ, Davis CM, Lindsay RC, Powell FL, Dan Y, Dickerman AW, Oosumi T, Shulaev V (2012) Transposon tagging in diploid strawberry. Plant Biotechnol J 10:985-994

Vollbrecht E, Duvick J, Schares JP, Ahern KR, Deewatthanawong P, Xu L, Conrad LJ, Kikuchi K, Kubinec TA, Hall BD, Weeks R, Unger-Wallace E, Muszynski M, Brendel VP, Brutnell TP (2010) Genome-wide distribution of transposed *Dissociation* elements in maize. Plant Cell 22:1667-1685

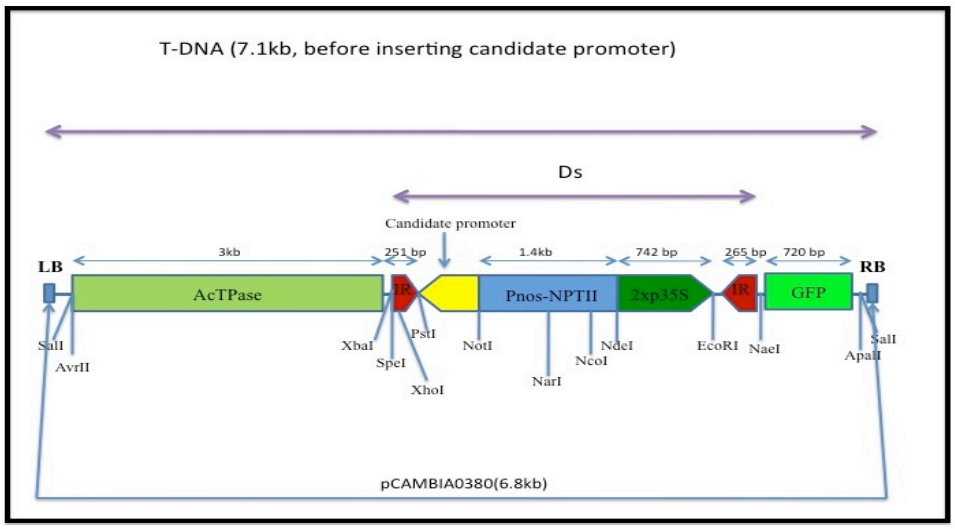


Figure 11 Schematic map of the tissue specific *pAcDs* constructs (*pAcDs-AtSCP* and *pAcDs-AmDEFH125*). LB, left border; RB, right border; *Ds*, *dissociator*; AcTPase, transposase gene; GFP, green fluorescent protein; IR, inverted repeat; 2xp35S, two times CaMV 35S basic promoter; Pnos, pnos promoter; *NPTII*, neomycin phosphotransferase II gene; Yellow arrow showed direction of the tissue-specific promoter, AtSCP or AmDEFH125. T-DNA was 7.1 kb without tissue-specific promoter. The whole cassette was inserted in vector *pCAMBIA0380*.

Table 5 Ploidy estimated by flow cytometry and T-DNA composition determined by multiplex PCR of 19 independent transgenic potato lines after transformation with either of two new constructs, *pAcDs-AtSCP* or *pAcDs-AmDEFH125*, carrying gametophyte-specific promoters to drive transposase.

pAcDs-AtSCP	**Ploidy**	**Multiplex PCR**
AtSCP-5C-1	2x	FDS, *NPTII*
AtSCP-B11	2x	FDS, EDS, *NPTII*
AtSCP-5D-1	2x	EDS
AtSCP-5F-1	4x	FDS, *NPTII*
AtSCP-P4	4x	FDS, *NPTII*
AtSCP-P11	2x	FDS, *NPTII*
pAcDs-AmDEFH125	**Ploidy**	**Multiplex PCR**
AmDEFH125-8	2x	FDS, *NPTII*
AmDEFH125-104	4x	FDS, *NPTII*
AmDEFH125-107	2x	FDS, *NPTII*
AmDEFH125-112	4x	FDS, *NPTII*
AmDEFH125-140	2x	FDS, *NPTII*
AmDEFH125-182	4x	FDS, *NPTII*
AmDEFH125-199	2x	EDS, FDS, *NPTII*
AmDEFH125-204	4x	FDS, *NPTII*
AmDEFH125-217	4x	FDS, *NPTII*
AmDEFH125-265	2x	FDS, *NPTII*
AmDEFH125-283	4x	FDS, *NPTII*
AmDEFH125-297	4x	FDS, *NPTII*
AmDEFH125-322	2x	FDS, *NPTII*

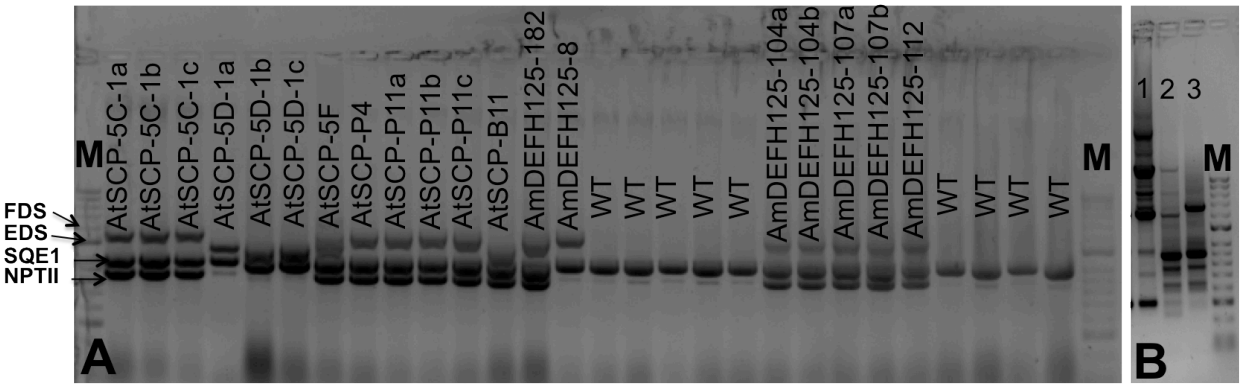
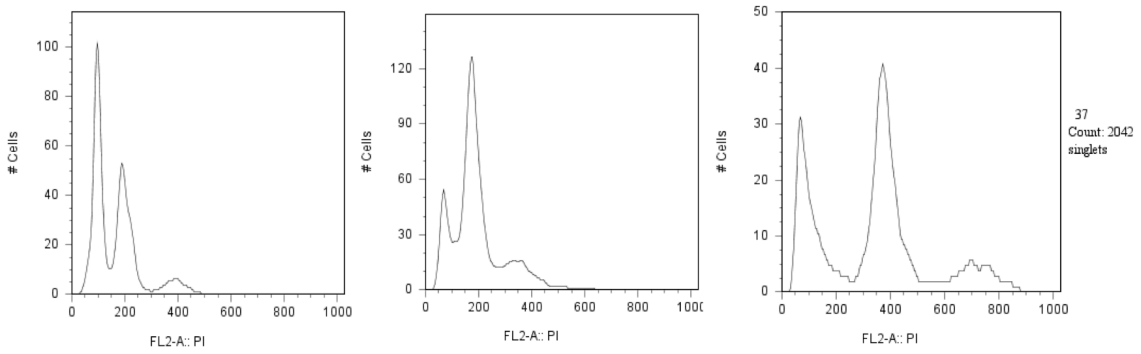


Figure 12 Multiplex PCR screening for transgenic potato (A) and HiTAIL-PCR for amplifying flanking sequences from launch pads (B). (A) Multiplex PCR screening of 28 potato plants (lanes 2-29) after transformation. The PCR contained primers to amplify empty donor site (EDS, 1 kb), full donor site (FDS, 1.1 kb), *SQE1* (800 bp) and *NPTII* (700 bp). The a, b, c at the end of each ID represented different shoots from the same callus. WT, wild type escapes. 19 of the 28 plants were transgenic, with four of them showing early transposition (AtSCP-5D-1a, b, c, AtSCP-B11) and nine were wild types. Lanes 1, 2 and 3 in (B) were HiTAIL-PCR results from AtSCP-B11 and AtSCP-P11a/b using RBTAIL primer sets (RB0a, RB1a and RB2a). Gel image was inverted for better quality. M, HyperLadder™ II.



Monoploid Diploid Tetraploid

Figure 13 Flow cytometry results from two transgenic potatoes to determine their ploidy. From left to right: monoploid potato control with 1x, 2x and 4x peaks, diploid transgenic AtSCP-5C-1 with 2x and 4x peaks (the indistinct peak to the far left is debris) and tetraploid potato AmDEFH125-104 with 4x and 8x peaks (the left most peak is debris).

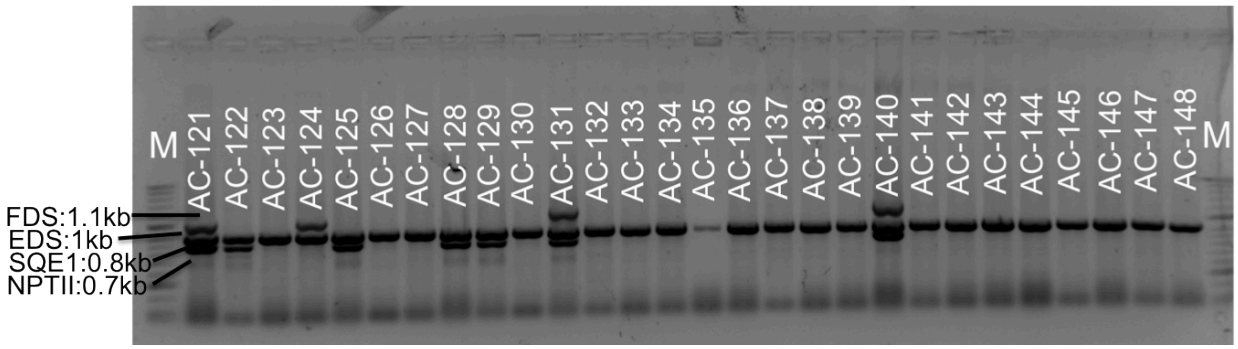


Figure 14 Multiplex PCR screening of 28 potato plants (AC-121 to AC-148) derived from anther culture. Plants were from AmDEFH125-104 and AtSCP-5C-1. The PCR contained primers to amplify empty donor site (EDS, 1 kb), full donor site (FDS, 1.1 kb), *SQE1* (800 bp) and *NPTII* (700 bp). Five transposants were found (AC-121, 122, 125, 128, 129), two were launchpads (AC-131, 140) and the rest were wild types. Gel image was inverted for better quality. M, HyperLadder^{TM} II.

Figure 15 EDS amplifications from plants carrying *Ac* after *Ds* excision. AC-6, 7, 20, 23 were anther culture derived potatoes from launch pad AmDEFH125-104; AC-46 was anther culture derived potato from launch pad AtSCP-5C. T_0-5D-1 (AtSCP-5D-1), T_0-199 (AmDEFH125-199) and T_0-B11 (AtSCP-B11) were three T_0 potatoes that showed early transposition. NEW ACDS EDS was the plasmid sequence without *Ds* part. Black arrow indicates sequences from transposase end or GFP end. Nucleotides with black background were same sequences. Nucleotides with white background were different from plasmid due to deletion, insertion or substitution. Dashes in picture indicate nucleotide deletions.

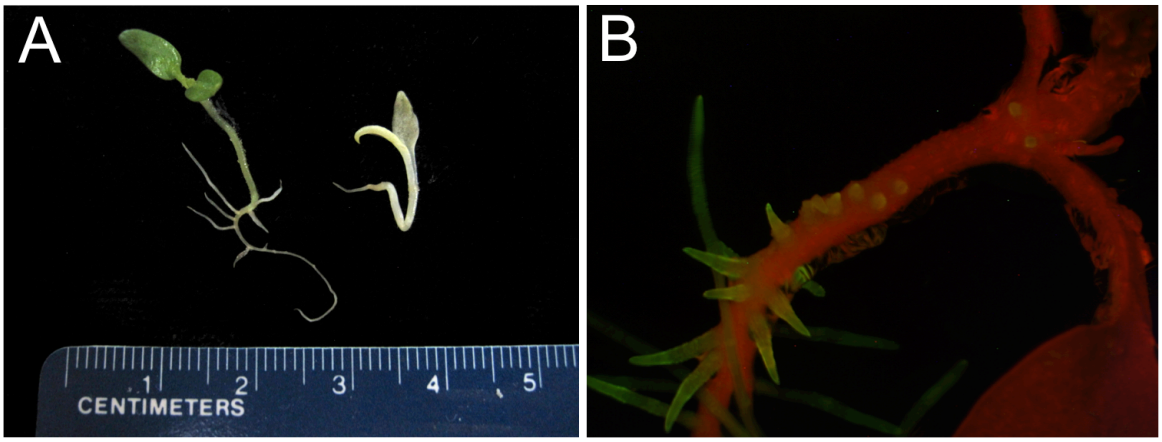


Figure 16 Screening of T_1 potatoes for putative transposants. A) Kanamycin resistant (left) and sensitive (right) seedlings 2 weeks after germination in liquid B5 basal medium with 50 mg/L kanamycin. Wild type seedling had pale cotyledon and short roots B) GFP expression in elongating roots from kanamycin resistant seedling.

Table 6 Screening of transposants from anther culture derived potatoes from two lines, AtSCP-5C-1 and AmDEFH125-104, and the HiTAIL-PCR results of two putative transposants, AC-9/10, AC-23. Reinsertion site of AC-27 was in the middle of the vector sequences, indicating local transposition into vector backbone that must have integrated into the potato genome during transformation.

A)

ID	FDS+*Ds*	EDS+*Ds*	*Ds*	EDS	WT
AmDEFH125-104	5 (11%)	5 (11%)	23 (50%)	1 (2%)	12 (26%)
AtSCP-5C-1	6 (8%)	1 (1%)	3 (4%)	1 (1%)	66 (86%)

B)

ID	LG	Genomic region	Gene annotation
AC-9, 10	chr05:43249504	Within gene PGSC0003DMG400027187	Proline transporter 3
AC-23	chr03:42196884	Promoter region of PGSC0003DMG400000639	Delta 9 desaturase
AC-27	NA	Into vector sequence	NA
Other ACs	NA	Adjacent to transposase gene	NA

Table 7 HiTAIL-PCR primer and location on linkage group (LG) of potato T_0 launch pads carrying A) *pAcDs-AtSCP* and B) *pAcDs-AmDEFH125* construct, as well as the genomic loci of insertion in genome browser.

A)

ID	TAIL primer	LG	Genomic loci	Gene annotation
AtSCP-P4	RB2a	chr09:48128599	Within gene PGSC0003DMT400044363	AT4G14385 protein
AtSCP-P11	RB2a	chr08:40101235	Promoter of gene PGSC0003DMT400010087	Pentatricopeptide repeat-containing protein
AtSCP-5D-1	RB2a	chr07:33805922	Within gene PGSC0003DMT400054002	Bem46
AtSCP-5F	RB2a	chr03:35757438	Intergenic	NA
AtSCP-5C-1	RB2a	Backbone	NA	NA
AtSCP-B11	RB2a	Backbone	NA	NA

B)

ID	TAIL primer	LG	Genomic loci	Gene annotation
AmDEFH125-8	RB2a	chr06:42636220	Intergenic	NA
AmDEFH125-104	RB2a	chr03:42409905	Intergenic	NA
AmDEFH125-112	RB2a	chr07:1911364	Intergenic	NA
AmDEFH125-182	RB2a	chr08:37227771	Intergenic	NA
AmDEFH125-204	RB2a	chr06:40761463	Promoter of gene PGSC0003DMT400068373	Cytochrome P450
AmDEFH125-283	RB2a	chr10:5306807	Within gene PGSC0003DMT400049950	Bypass2
AmDEFH125-265	RB2a	chr12:5434357	Intergenic	NA
AmDEFH125-217	RB2a	NA	NA	NA
AmDEFH125-297	RB2a	NA	NA	NA
AmDEFH125-322	RB2a	NA	NA	NA
AmDEFH125-140	RB2a	NA	NA	NA
AmDEFH125-199	RB2a	Backbone	NA	NA
AmDEFH125-107	RB2a	Backbone	NA	NA

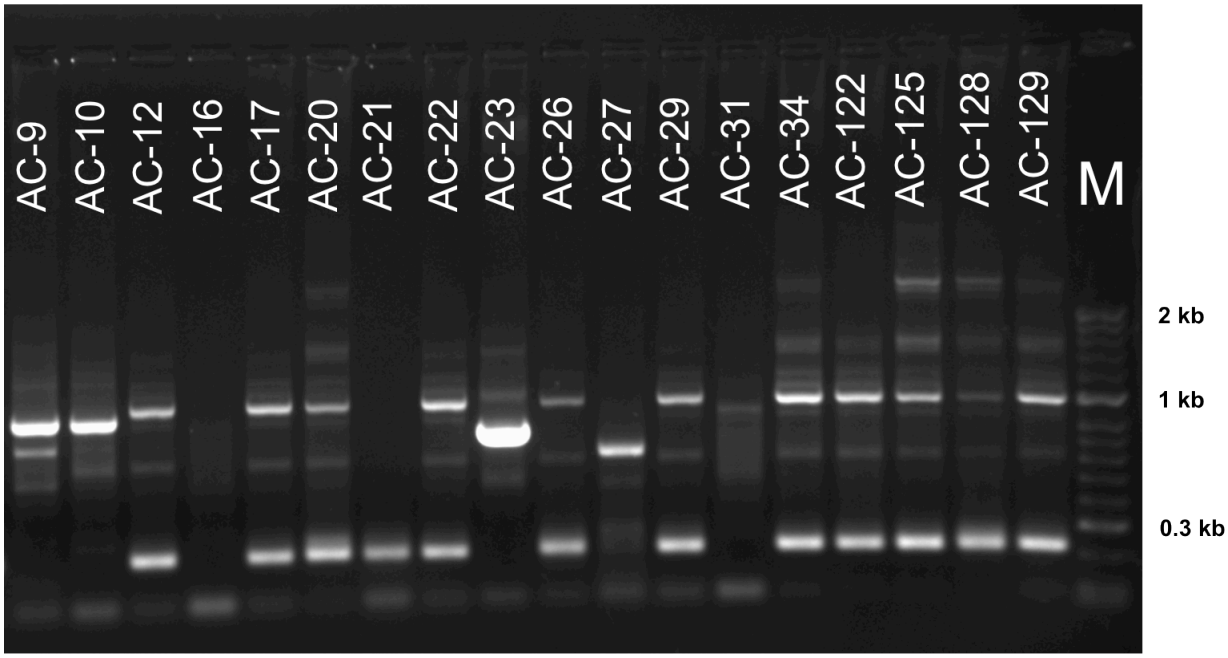


Figure 17 HiTAIL-PCR of 18 candidate transposants derived from anther culture using *Ds*5 primers sets (*Ds*5-1, *Ds*5-2, *Ds*5-3). The PCR amplified flanking sequences from inverted repeat (IR) using *Ds*5 primer sets. Unique bands were found in AC-9, 10, 23 and 27. Two common bands were found in other samples. M, HypperLadder^{TM} II.

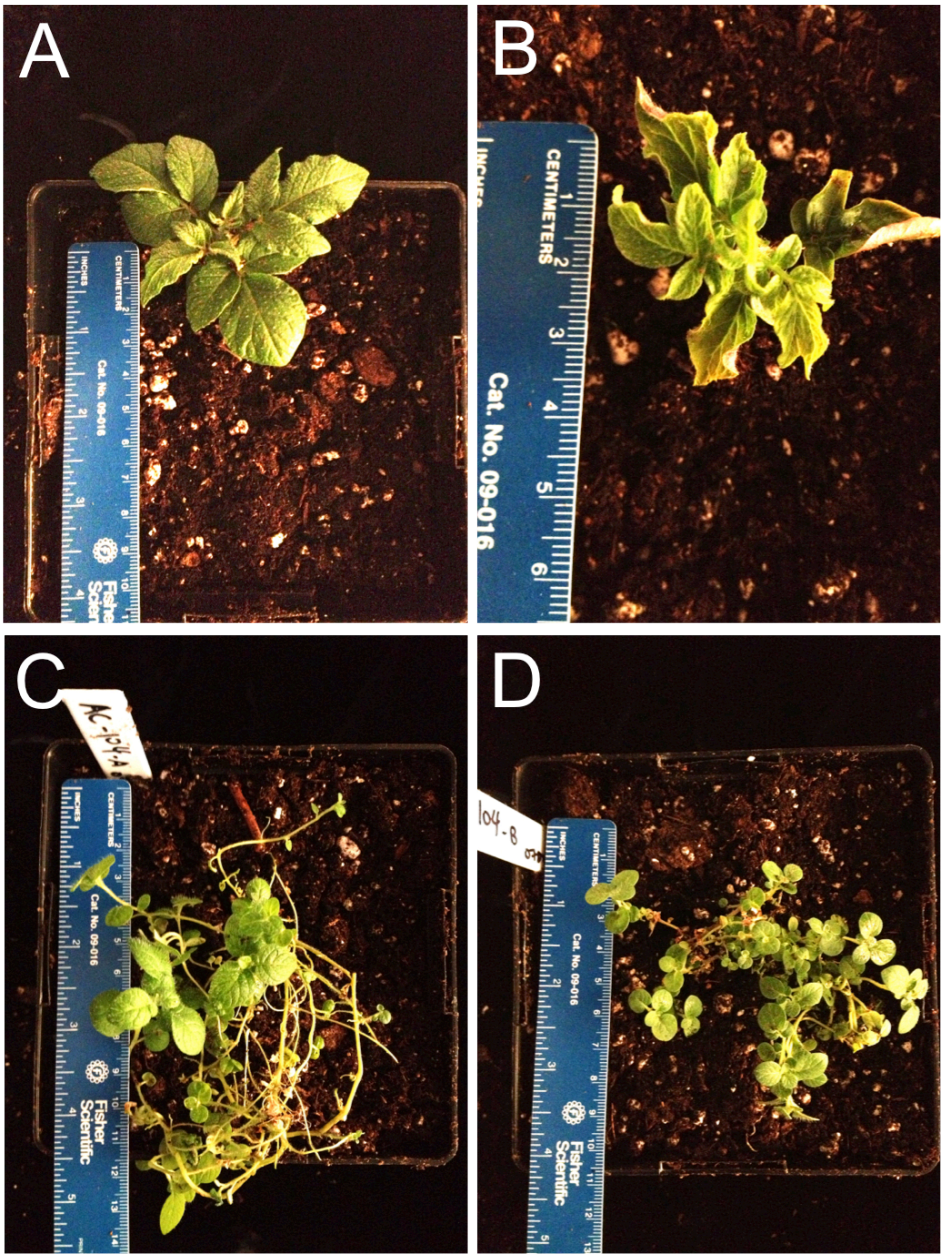


Figure 18 Mutant pictures from screening of transposants during this study. (A) Wild type T_1 plant (B) Curly leaf T_1 plant (ID: 5C-101) derived from crosses between AtSCP-5C-1 and RH. (C) Anther culture derived candidate transposant with *Ds* only from line AmDEFH125-104. (D) Anther culture derived candidate transposant with EDS+*Ds*.

Supplemental material

Codon optimization of transposase

```
Amino Acid: Met Ala...Pro...Lys...Ser...Glu...Leu...Glu...Gly...Val
Codon:      ATG GCT...CCa...AAa...TCt...GAa...CTc...GAa...GGt...GTt
Position:    1   2    25   36   39   78  130  147  235  274

Amino Acid: Ala...Ser...Asn...Ala...Pro...Pro...Val...Ala...Ser...Glu
Codon:      GCt...TCt...AAc...GCt...CCa...CCa...GTt...GCt...TCt...Gag
Position:   324  355  363  478  483  575  606  619  636  640

Amino Acid: Arg...Glu...His
Codon:      AGg...GAa...CAc
Position:   653  667  682
```

Codon optimization of *NPTII*

```
Amino Acid: Met Ile...Gly...Leu...Ala...Arg...Thr...Gly
Codon:      ATG ATT...GGa...CTt...GCt...CGt...ACt...GGa
Position:    1   2    43   60   78  177  187  210
```

Supplemental Figure 2 Codon optimization of transposase and *NPTII* gene. Upper case letters in codon sequences represent same nucleotide whereas lowercase letter represent modified sequences. Numbers in transposase were counted as amino acid position in open reading frame b.

Supplemental Table 4 Primers used in this study.

Primer name	Primer sequence (5'-3')	Purpose
Not1.AtLBD27F	CACCGCGGCCGCAGTTCCACATGTGCTCCGGTAATC	Cloning
Pst1.AtLBD27R	ACAACTGCAGTCGATGGGAACAGAACAAGAGACTC	Cloning
Not1.FvDEFH125F	CACCGCGGCCGCGCGCATCTACAAGCACTAAATCTC	Cloning
Pst1.AmDEFH125R	ACAACTGCAGCAAATGAAATTTCTAGGGTTTCTTCTC TG	Cloning
TPaseR1	GGTGAAATGCTGCCATAC	Multiplex PCR
EGFPR1	TTGTACAGCTCGTCCATG	Multiplex PCR
p35SF1	ACGCACAATCCCACTATC	Multiplex PCR
NPTIIF2	CGGCTATGACTGGGCACAACAG	Multiplex PCR
NPTIIR2	GGCGATACCGTAAAGCACGAGG	Multiplex PCR
SQEF	TGGGGTTCGTTGCAGTTTTC	Multiplex PCR
SQER	CGCGTTGAGCATCAATTTTCTC	Multiplex PCR
RB-0a	GGCAATAAAGTTTCTTAAGATTGAATCCTGT	TAIL-PCR
RB-1a	ACGATGGACTCCAGTCCGGCCTGTTGCCGGTCTTGCG ATGATTATCA	TAIL-PCR
RB-2a	GTAATGCATGACGTTATTTATGAGATGGGTT	TAIL-PCR
LB-0	GTAGTTCCCAGATAAGGGAATTAGGGT	TAIL-PCR
LB-1	ACGATGGACTCCAGTCCGGCCCCTATAGGGTTTCGCT CATGTGTTGA	TAIL-PCR
Ds3-1	ACCCGACCGGATCGTATCGGT	TAIL-PCR
Ds3-2	ACGATGGACTCCAGTCCGGCCCGATTACCGTATTTAT CCCGTTC	TAIL-PCR
Ds3-3	GTATTTATCCCGTTCGTTTTCGT	TAIL-PCR
Ds5-1	ACGGTCGGGAAACTAGCTCTAC	TAIL-PCR
Ds5-2	ACGATGGACTCCAGTCCGGCCCCGTTTTGTATATCCC GTTTCCGT	TAIL-PCR
Ds5-3	TACCTCGGGTTCGAAATCGAT	TAIL-PCR
LAD1-1	ACGATGGACTCCAGAGCGGCCGCVNVNNNGGAA	TAIL-PCR
LAD1-2	ACGATGGACTCCAGAGCGGCCGCBNBNNNGGTT	TAIL-PCR
LAD1-3	ACGATGGACTCCAGAGCGGCCGCVVNVNNNCCAA	TAIL-PCR
LAD1-4	ACGATGGACTCCAGAGCGGCCGCBDNBNNNCGGT	TAIL-PCR
AC1	ACGATGGACTCCAGAG	TAIL-PCR

4. CHARACTERIZATION OF A STRAWBERRY MUTANT OF THE SUCROSE TRANSPORTER-2 GENE

Abstract

Sucrose is one of the final products of photosynthesis in higher plants. Sugar transportation from source leaf into sink leaf and other organs is important in maintaining normal functions of the plant life cycle. Sugar transport in strawberry would be expected to affect fruit development and flavor. A mutant of sucrose transporter-2 gene (hybrid gene model gene27493, *FvSUC2*) in *Fragaria vesca*, the diploid strawberry, which showed obvious stunted growth, as well as enhanced anthocyanin pigmentation was recovered in the progeny of a plant that had been transformed with an *Activator/Dissociator* construct. The T-DNA insertion was in the 3'UTR region of the sucrose transporter-2 gene; however, a single bp deletion in exon 4 of the gene was also discovered, thus confounding the specific cause of the phenotype. Metabolite profiling of leaves of the mutant using GCMS revealed significantly greater concentrations of various carbohydrate compounds including fructose, glucose, sucrose compared to wild type plants, indicating the function of this gene in loading sugars into vascular tissues. Phylogenic analysis placed this gene within group 2 in this gene family where similar phenotypes have been observed in knockout mutations of Arabidopsis and tomato.

Key words: *Fragaria vesca*, sucrose transporter-2 gene, T-DNA insertional mutant, metabolite profiling

Abbreviations- *AcDs*, *Activator/Dissociator*; EDS, empty donor site; FDS, Full donor site; GCMS, Gas chromatography–mass spectrometry; qRT-PCR, quantitative real-time PCR;

Introduction

Strawberry is an important fruit crop composed of high content of different nutrients and vitamins (Shulaev et al. 2008; Shulaev et al. 2011). The United States is one of the world's leading producers of strawberries, especially in Florida and California. Many previous studies of strawberry have been focused on increasing disease and pest resistance (Davik and Honne 2005; Martínez Zamora et al. 2008; Vellicce et al. 2006), improving nutritional values (Agius et al. 2003), and understanding fruit developmental processes (Jia et al. 2013; Jia et al. 2013; Salvatierra et al. 2013; Thill et al. 2013). The commercial strawberry, *Fragaria* × *ananassa*, is an octoploid species with complex genome content. The diploid strawberry, *Fragaria vesca*, however, is a diploid species with many good features to provide a nice model fruit plant system, including short life cycle, easy transformation, prolific seed production, and easy cultivation (Shulaev et al. 2011). Compared to the red fruit of *Fragaria* × *ananassa*, *Fragaria vesca* has white or red fruit in different accessions. With the genome sequence of *Fragaria vesca* published, many candidate genes that are related to flowering, flavor and fruit ripening have been annotated (Shulaev et al. 2011). In order to characterize genes of interest, Veilleux et al. (2012) generated several populations of T-DNA insertional or transposon tagged mutants. Within these populations, many mutants with remarkable phenotype have been discovered, including a sucrose transporter-2 gene mutant, providing a good resource for forward genetics.

The sucrose transporter gene family belongs to the major facilitator superfamily of genes. Many of its members have been well-characterized in Arabidopsis (Williams et al. 2000), potato (Chincinska et al. 2008; Kuhn et al. 2003), tomato (Hackel et al. 2006; Weise et al. 2008), strawberry (Jia et al. 2013), rose (Henry et al. 2011), etc. In Arabidopsis, there are nine members that have been identified in the sucrose transporter gene family, *AtSUC1* to *AtSUC9* (Williams et al. 2000). The sucrose transporter genes have specific affinity for different substrates and function either in long- or short-distance transport, loading/unloading of sugars, or pollen development (Sauer 2007). Thus these gene members have been categorized into four groups based on their functionality and their evolutionary distance (Sauer 2007). The study of promoter regions of these genes revealed a few motifs in the *cis*-element that might be important in regulating expression

(Schneidereit et al. 2008). The transgene expression of β-glucuronidase (GUS) in strawberry under the control of the *AtSUC2* promoter showed strong GUS expression in phloem of leaf tissue (Zhao et al. 2004) indicating a similar *cis*-element or motif existed in both species.

In a study of all seven candidate sucrose transporters genes in *Fragaria* × *ananassa*, Jia et al. (2013) found that the seven genes had different expression levels in different tissues and that *FaSUT1* played a key role in strawberry fruit ripening compared to the other homologues (Jia et al. 2013). The Arabidopsis *AtSUC2* gene was found to be important in phloem loading but not in long-distance transport, and its substrate of transport was not limited to sucrose but also included arbutin, ά-paranitrophenylglucoside (Chandran et al. 2003; Srivastava et al. 2008). The same scenario occurred in other species as well. The three homologous genes in grape vine, which was closely related to strawberry, were also found to have totally different expression patterns (Davies et al. 1999). It was proposed that mutation in genes responsible for phloem loading of this family might be more severe than mutations in other members of this gene family. It has also been found that sucrose transporter-2 gene in rose was involved in bud burst (Henry et al. 2011).

In a study of a tomato sucrose transporter, it was found that both the introns and 3'UTR contributed to gene expression (Weise et al. 2008). Research in cancer and other projects has revealed the function of 3'UTR in alternative splicing (Mayr and Bartel 2009). Many databases have been established containing publically available 3'UTR sequences and motifs in plant and animal kingdoms (Grillo et al. 2010). These sequences provide good resources for understanding the roles of 3'UTR on transcriptional and translational levels. In this study we characterized a knockout mutant in *Fragaria vesca* with sugar transport deficiency due to T-DNA insertion in the 3'UTR region. Sugar compounds including fructose, glucose and sucrose were significantly more abundant in leaves of mutant compared with wild type plants, suggesting the function of *FvSUC2* in loading sugars into vascular tissue, as well as in transport.

Materials and methods

Plant material and growth conditions

A single T-DNA mutant, LP119 and its progeny, of the diploid strawberry, *Fragaria vesca* "Hawaii-4" (PI 551572) was used for in this study. The mutant had been classified as a putative launch pad after transformation with an *activator/dissociator* T-DNA construct, *pAcDs*, as described in Veilleux et al. (2012). T_1 and T_2 seeds were both collected after self-pollination of T_0 and T_1 plants, respectively. Strawberry seeds were germinated in liquid B5 medium (B5 basal medium, 2% sucrose, pH 5.8) and germinated seeds were transferred to closed-environment growth chambers (Conviron model ATC60) at day neutral conditions (12 h photoperiod, 22°C/16°C thermoperiod, with light intensity of 175 μm/m^2/s). Plants were then moved to the greenhouse after 4 months for fruit set and seed collection. For T_2 analysis, wild type PI 551572 strawberries were treated using the same conditions as for the mutants.

DNA extraction and PCR characterization

A multiplex PCR comprising primers to amplify empty donor site (EDS), full donor site (FDS), housekeeping gene *FvAAT3* (alcohol acyltransferase 3), as well as the kanamycin resistance gene *NPTII*, were used to genotype the T_0, T_1 and T_2 plants (Veilleux et al. 2012). HiTAIL-PCR was performed using primers designed from the transposase gene or inverted repeats (IR) to amplify flanking sequences from launch pad or *Ds*. Zygosity test was performed using primers from transposase (newtpaseRV-2a) and primers from flanking genomic regions (338ZygF/R). DNA extraction and PCR programs were as described in (Veilleux et al. 2012).

Southern blot

Genomic DNA was extracted from young leaf tissue and digested with *HindIII*. Digested DNA was hybridized with radioactively labeled probe designed from FDS to detect the copy number of T-DNA (*pAcDs*) insertions in the T_0 plant (Veilleux et al. 2012).

RNA extraction and semi-qRT PCR

RNA was extracted from young, fully expanded leaf tissue of growth chamber plants. RNA extraction was as described by (Yu et al. 2012). Primers for both gene Fv27493 (*FvSUC2*) and Fv27492 were designed from two flanking exons with one intron. *FvActin* was used as control primer to test for successful extraction. Total RNA was treated with *DNAseI* using Ambion TURBO DNA-free^{TM} kit (AM1907), and then 1 ug of clean RNA was reverse transcribed to cDNA using Thermo DyNAmo SYBR Green 2-Step qRT-PCR kit (# F-430L); 100 ng of synthesized cDNA was used as template for semi-quantitative RT-PCR reactions.

Sequencing of *FvSUC2* gene in mutant

Primers (FvSUC2F/Fv27493R and FvSUC2R/Fv27493F, Supplemental Table 5) designed from *FvSUC2* were used to amplify the gene sequence by PCR (iProof^{TM} High-Fidelity master mix, Bio-rad). PCR products were loaded in 0.8% agarose gel and extracted bands were purified using QIAquick gel extraction kits (catalog 28704). PCR was repeated and sequenced three times to confirm sequence accuracy. The quality of sequences was examined by BioEdit software. Sequence analysis was conducted using Lasergene 10 software.

GC-MS analysis

Source leaf tissue was collected from greenhouse plants and frozen in liquid nitrogen immediately before lyophilization for 18 h in FreeZone freeze drier. Dried (1 mg) tissue was weighed and added to extraction buffer containing 200 uL chloroform and 200 uL 10 mM HCl; 2uL of 5mM ribitol were added at the same time to be used as internal standard. Samples were vortexed vigorously for 1 min and centrifuged at 13,000 rpm for 2 min after which 50 uL of supernatant were transferred to new glass vials with inserts. Samples were dried under N_2 at 50°C. Then 25 uL of methoxyamine HCl in pyridine were added to each sample before a 2 h incubation at 50°C. Then 25 uL of MSTFA (N-methyl-N-trimethylsilyl-trifluoroacetamide) + 1% (v/v) TMCS (trimethylchlorosilane) were added and incubated for 30 min at 50°C. Two injections, 1 uL and 0.1 uL, from the same sample were injected and separated in the Agilent 7890A series GC equipped with a

DB-5MS-DS column and analyzed on an Agilent 5975 C series single quadrupole mass spectrometer (MS). Automated Mass Spectrometry Deconvolution and Identification System (AMDIS, NIST) was used to deconvolute signals of the coeluting compounds. Using the Enhanced Mass Selective Detector ChemStation software (Agilent Technologies) in combination with the custom-built spectral and retention time library, the spectral NIST library (National Institute of Standards and Technology, Gaithersburg, MD), we identified different compound. The identities of metabolites and quality of integration were curated manually on individual basis by using the QEdit function of the ChemStation software after automated peak area integration. Relative areas of the internal standards were used to correct for recovery in quantitation of each metabolite and all samples were standardized with respect to the dry weight of the tissue used for extractions. Signals from the 0.1 uL injection were used exclusively for detection of fructose, glucose, myo-inositol and sucrose. Signal from the 1 uL injection was used to detect other low yield polar metabolites. P-value of sugar compound was analyzed using student t-test. Principle component analysis was performed in Jmp10 software.

Results

Mutant screen

In the population that was generated after transformation with the *pAcDs* construct (Fig. 19), T_0 338 (LP119) was a single T-DNA insertion mutant with its insertion site located at chr02: 13951314 confirmed by HiTAIL-PCR (Fig. 20). After transformation, single insertion T_0 plants could carry the entire *AcDs* construct and thus be launch pads with a full donor site (FDS), carry only the *Ds* element and thus be somatic transposants or carry an empty donor site (EDS), with or without a separate *Ds* element, and thus be insertional mutants incapable of further transposition because the promoter would have been separated from the transposase gene. Multiplex PCR of transgenic plant LP119 confirmed the presence of FDS and no EDS; thus it was classified as a putative launch pad destined for T_1 analysis. A BLAST search of the nucleotide sequence obtained by HiTAIL-PCR using nested primers embedded in the transposase gene of the *AcDs* construct results against the strawberry genome browser (https://strawberry.plantandfood.co.nz/cgi-bin/gbrowse/strawberry_genome/) revealed the insertion to be 226 bp downstream of the

stop codon of Fv27493 (hybrid gene model, sucrose transporter-2, *FvSUC2*), expected to be within the 3' UTR region (Fig. 21). The insertion was also 1,530 bp upstream of the start codon of gene model Fv27492, a zinc-finger CCCH domain containing protein (Fig. 21). Plants were allowed to self-pollinate and T_1 seed was collected.

Among six T_1 plants grown to screen for transposition, segregation was observed for both FDS and EDS (Fig. 20C). T_1 plant A5 was observed to have increased anthocyanin pigmentation in the runners and petioles. PCR characterization of this mutant revealed EDS, FDS and *NPTII* in its genome (Fig. 20C), implying that it was homozygous for the T-DNA at the insertion site, with both and intact and a transposed copy as well as a likely transposed *Ds* element reinserted elsewhere in the genome. No obvious phenotype was observed in the T_0 or other T_1 plants screened.

T_2 zygosity test

A T_2 seedling generation was grown to observe segregation of the various T-DNA components. All 21 T_2 plants that were screened had *NPTII*, confirming presence of the *Ds* element; however, based on segregation of FDS and EDS, we could recognize three classes of progeny. Six plants had FDS and *NPTII*, seven plants have EDS and *NPTII*, and eight plants had EDS, FDS and *NPTII* (Fig. 22). No wild type plants were found (Table 8). Primers designed for zygosity revealed all 21 T_2 plants to be homozygous for the Tpase/*FvSUC2* region (Fig. 23).

Extra *Ds* reinsertion in genome

Since EDS was found in several plants, there was a second *Ds* reinsertion site in the strawberry genome besides launch pad. In order to better understand the effects of the extra *Ds* reinsertion in the genome, we performed HiTAIL-PCR on four mutants, *fvsuc2-2*, *fvsuc2-4*, *fvsuc2-5*, *fvsuc2-28*, using *Ds*5 primer sets, and found that the *Ds* insertion fell into two positions (Supplemental Fig. 3, Supplemental Table 6). For *fvsuc2-2/5/28*, the *Ds* reinserted in chr02: 31941485, intron region of gene16106, a candidate gene with unknown function, UPF0558 protein C1orf156; and for *fvsuc2-4*, the reinsertion site was chr06: 36637124, an intergenic region in genome.

Among all 21 T_2 plants, regardless of the of an extra *Ds* insertion in the genome (i.e., EDS+*NPTII*), all plants exhibited the same phenotype of stunted growth with increased levels of anthocyanin in leaves and petioles (Fig. 24, supplemental Fig. 4). The HiTAIL-PCR results also confirmed two reinsertion sites of *Ds* in different chromosomes from four plants, either within a gene of unknown function or in intergenic region. Together, these data indicated that the phenotype was due to the original T-DNA insertion rather than *Ds* reinsertion.

To better understand the function of *FvSUC2*, we selected three T_2 plants (*fvsuc2-4, fvsuc2-5, fvsuc2-28*), as well as four wild type plants for GC-MS analysis of carbohydrate concentration. Among these three mutants, *fvsuc2-4* had EDS, FDS, and *NPTII*, whereas *fvsuc2-5* and *fvsuc2-28* had EDS and *NPTII* (Fig. 22). The three mutants had two different *Ds* reinsertion sites.

Phenotypic analysis

Plant height, total area of rosette leaves, fruit number and maximum runner length measured in the greenhouse five months after germination were significantly less ($p<0.01$) for the mutant compared to wild type (Fig. 25, Supplemental Fig. 4). These phenotypic data for strawberry resembled that reported for the Arabidopsis *AtSUC2* knockout mutant (Wippel and Sauer 2012). The smaller leaves, smaller fruit size and increased anthocyanin content all suggested stress due to sucrose accumulation in the source leaf.

Semi-quantitative RT-PCR

Leaf tissue from mutant and wild type plants was used for RNA extraction and semi RT-PCR. No amplification of gene *FvSUC2* was found in any of the four mutants, while strong bands were found in two wild type plants. Amplification of Fv27492 situated about 1.5 kb from *AcDs* insertion was observed in both mutant and wild type plants (Fig. 26). This suggests that the T-DNA insertion in the 3'UTR region of *FvSUC2* had inhibited expression of this gene.

GC-MS analysis

Concentrations of 25 compounds were identified after GC-MS analysis, including different forms of carbohydrates (monosaccharide, disaccharide, trisaccharide). Principle component analysis separated mutant from wild type samples (Fig. 27). The concentrations of monosaccharides and disaccharides (fructose, glucose, sucrose) were greater in mutant compared to wild type (Fig. 28A), whereas trisaccharide concentrations (galactinol, raffinose) were greater in wild types (fig. 28B). These data indicate that sucrose may not be the only substrate for *FvSUC2*. Because galactinol and raffinose have been previously reported to be involved in plant adjustment to stress (Ito et al. 2005), greater concentrations of these trisaccharides could be expected in mutants compared to wild type plants. Myo-inositol, another carbohydrate which is an important signaling compound (Datta 2005), was not significantly different between the two groups.

Carbohydrate concentrations among the three mutants with different *Ds* reinsertion sites were not significantly different, which supported our theory that the original T-DNA insertion being the cause of the mutant phenotype.

Sequencing of *FvSUC2* gene in mutant revealed a one-nucleotide deletion in coding region

Primers designed to clone the whole gene were used to amplify *FvSUC2* from the mutant. Sequences were used to align with *FvSUC2* sequences from genome browser, and a one-nucleotide deletion in the 4th exon of the mutant was found, which could have caused a frame shift and early truncation of gene translation (Fig. 29). It is possible that this frame shift and the T-DNA insertion together lead to the inhibited expression of *FvSUC2*.

mRNA and 3'UTR sequences alignment of *FvSUC2* with other publically available homologs

The mRNA sequence retrieved from the gene model was used to align with *AtSUC2* (Genbank accession X79702) and *RhSUC2* (Genbank accession HQ403679) that were characterized from Arabidopsis and rose (Fig. 30A). Sequences were 70% similar at the amino acid level between *AtSUC2* and *FvSUC2*, and 91% similar between *RhSUC2* and *FvSUC2*. Twelve conserved domains (I-XII in Fig. 30A) were identified that might be

involved in membrane spanning (Davies et al. 1999). The 3'UTR sequences of *RhSUC2* and *FvSUC2* were also aligned and the sequences were 86% similar (Fig. 30B), suggesting highly conserved domains during evolution.

Phylogenetic study of *FvSUC2*

The mRNA sequences of *FvSUC2* were aligned in Genbank database and analyzed using neighbor-joining method. The *FvSUC2* gene was closely related to the characterized *RhSUC2* and *FaSUC4* proteins, as expected, suggesting it was involved in bud burst and fruit development. The phylogenetic study showed *FvSUC2*, *RhSUC2* and *AtSUC2* were in the same group (Fig. 31), as discussed above.

The group 2 sucrose transporters represent membrane-localized sucrose transporters from dicots (Sauer 2007). Their main functions are to load sucrose into phloem or sink tissues (Sauer 2007), so the accumulation of excessive sucrose in mutant source leaf could cause stress and lead to increased anthocyanin contents, as found in Arabidopsis and tomato mutants (Hackel et al. 2006; Wippel and Sauer 2012).

Discussion

Importance of sucrose transporters in plants

Sucrose is one of the final products of photosynthesis in higher plants and plays an important role in life cycle. Synthesized sucrose is transported from source leaf into sink leaf or fruits. Sucrose transporters belong to a large gene family, the major facilitator superfamily. Due to gene duplication, there tends to be multi members in the family in some dicot species, such as Arabidopsis.

There are nine members of sucrose transporters identified in Arabidopsis, *AtSUC1* through *AtSUC9*. Their function varies from phloem loading to unloading, or even in pollen development. They can be divided into several subfamilies (Sauer 2007). Knockout mutants in these genes do not always show phenotype, suggesting different importance.

Studies of strawberry as a non-climacteric fruit ripening have focused on abscisic acid (ABA) rather than ethylene, and sucrose has been suggested as an important trigger for

ABA in strawberry (Li et al. 2011). Thus success transportation of sucrose from source leaf into fruits would be important in fruit flavor and size, which is supported by our results (Fig. 24, Supplemental Fig. 4).

Mutant discovery

In our previous study, a large collection of strawberry insertional mutants was established. This includes activation tagged (*Ds* only) or T-DNA insertional mutants (*Ac/Ds* or *Ac*). The *fvsuc2* mutant was selected from this insertional population. The T_0 parent, which is hemizygous for T-DNA (FDS only), has no particular phenotype, while the T_1 mutant homozygous for T-DNA insertion (EDS and FDS) had the obvious phenotype. The sequencing result also showed a one-nucleotide deletion in the 4th exon, close to 3'UTR. So the actual reason leading to deactivation of *FvSUC2* could be the T-DNA insertion or by both T-DNA and the frame shift in coding region.

The T_2 plants derived from self-pollination of the mutant were all homozygote and showed the same phenotype as the T_1 parent. They all showed the same phenotype no matter with or without extra *Ds* reinsertion. And based on the similarity with mutant phenotype in Arabidopsis and tomato, they together provide strong evidence that the phenotype was due to the original T-DNA insertion rather than *Ds* reinsertion.

A second *Ds* insertion did not affect phenotype and sugar metabolites

HiTAIL-PCR results confirmed the insertion site of *Ds* in the genome for four mutants, *fvsuc2-2/4/5/28*, in which *fvsuc2-4* was different from the other three samples. But in the metabolite profiling using three samples, the sugar content of *fvsuc2-4* was not significantly different from the other two samples, *fvsuc2-5* and *fvsuc2-28,* while the phenotypes such as height, fruit set, leaf size, etc., were not significant either. These results confirmed the theory that the re-insertion of *Ds* was not the reason causing the phenotype or sugar contents in mutants compared to wild type strawberries.

Resemblance of sucrose transporter-2 mutants in different species

Sugar transportation from source leaf to sink leaf, stem or root is important for maintaining life functions; changes in transportation will be affecting pollen or fruit

development. The study of these transporters will provide useful information about metabolite engineering or new generation of clean energy (Chandran et al. 2003). Knockout plants in Arabidopsis and tomato showed similar phenotype, like dwarf, high concentration of anthocyanin, fewer fruits/seeds setup, and higher concentration of soluble sugars in source leaf (Supplemental Fig. 5) (Hackel et al. 2006; Wippel and Sauer 2012; Zhang et al. 2011). In the mutant *fvsuc2* that we discovered, we found highly resemblance both in phenotype and in metabolite profiling with *atsuc2* and *lesut2*. Based on phylogenetic study, *FvSUC2* and *AtSUC2* belong to the same group and their mRNA sequence is 67% similar. These members in group 2 were plasma membrane-localized sucrose transporters from dicots (Sauer 2007).

Interestingly, as shown in (Supplemental Fig. 5), in characterizing the Arabidopsis mutants, hemizygote plants show the same phenotype as wild types, but homozygote plants show severe phenotype as describe above (Wippel and Sauer 2012). The same happened to the strawberry mutant that the LP119 showed normal phenotype but the T_1 and T_2 homozygote plants showed severe phenotype.

Conserved *cis*-regulatory elements in promoter region

Function of gene expression is based on existence of *cis*-regulatory elements in promoter region. They determine when and where a gene expression would start. For sucrose transporter gene family, a few *cis*-regulatory elements have already been identified in tomato, Arabidopsis and apricot (Schneidereit et al. 2008). These domains were supposed to be putative binding site for transcription factors that determine when and where expression of the gene started. In scanning the predicted promoter region of *FvSUC2*, the same motifs were also found about 1.1 kb from starting codon (Supplemental Fig. 6). This provides strong evidence of similar function between *AtSUC2* and *FvSUC2*.

Conserved sequences in 3'UTR region between genes *RhSUC2* and *FvSUC2*

In comparing the 3'UTR sequences from *RhSUC2* and *FvSUC2*, the similarity is found to be 86%. The alignment between *AtSUC2* and *FvSUC2* or *RhSUC2* is not high (data not shown) suggesting it is not conserved between these 2 species. Arabidopsis plants with insertion in 3'UTR (T-DNA lines SALK_001331 and CS810256) were also grown under

long day conditions but no obvious phenotypes were found (data not shown). These data suggest the 3'UTR region in *FvSUC2* might be important in regulating gene expression. Study of tomato shows importance of introns and 3'UTR in regulating gene function (Weise et al. 2008).

Function of 3'UTR in gene regulation

Studies of the 3'UTR in gene regulation have been conducted in animals and plants (Tanguay and Gallie 1996; Wilkie et al. 2003). In a study of an Arabidopsis T-DNA insertional mutant, *rat5*, it was hypothesized that the T-DNA insertion in the 3'UTR of the *RAT5* histone H2A gene was the sole cause for the rat phenotype (Mysore et al. 2000). Subsequent complementation study confirmed the hypothesis and the result further suggested the importance of 3'UTR in maintaining stability and localization of mRNA to be detected.

Several other reports in studying other genes have found the T-DNA insertion in 3'UTR region disrupted gene expression or even caused severe phenotypes, suggesting their roles in regulation of gene transcription (Hu et al. 2009; Kim et al. 2011; Xiong et al. 2009; Zainal and Turnbull 2011). Our results here provided new resource for studying 3'UTR function in plant genomics.

Substrate specificity of *FvSUC2*

Plant sucrose transporters are divided into different phylogenetic groups and members belonging to different groups have different substrate specificity (Reinders et al. 2012). Studies in *Xenopus* proved such theories (Chandran et al. 2003). Sucrose, maltose, and other glucosides were all substrates in certain species.

Here in our study, the metabolite profiling experiment using GCMS suggested that sucrose might not be the only substrate *FvSUC2* carried into companion cells for transport. Other compounds such as fructose and glucose concentrations were also affected in mutants against wild type plants. Further experiments are needed in order to better understand which substrates were carried by *FvSUC2*.

The primers used in this study are shown in supplemental materials (Supplemental Table 5).

Acknowledgments

This work was supported by USDA⁄ NRI Award No. 2008-35300-04458, USDA⁄ NIFA Award Nos. VAW-2010-01562 and VAW-2009-04069, USDA Award 10-2361-04 (Biodesign and Bioprocessing Center), The Virginia Tobacco Indemnification Program, Virginia Agricultural Experiment Station Hatch Project 135816, the Virginia Tech Department of Horticulture and the Virginia Tech Molecular Plant Sciences Program. We thank Dr. Eva Collakova and Yihui Fang from Plant pathology, physiology and weed science department of Virginia Tech for their kind help in GCMS experiment and analysis.

References

Agius F, González-Lamothe R, Caballero JL, Munoz-Blanco J, Botella MA and Valpuesta V (2003) Engineering increased vitamin C levels in plants by overexpression of a D-galacturonic acid reductase. *Nat Biotechnol* **21**:177-181.

Chandran D, Reinders A and Ward JM (2003) Substrate specificity of the *Arabidopsis thaliana* sucrose transporter AtSUC2. *J Biol Chem* **278**:44320-44325.

Chincinska IA, Liesche J, Krugel U, Michalska J, Geigenberger P, Grimm B and Kuhn C (2008) Sucrose transporter StSUT4 from potato affects flowering, tuberization, and shade avoidance response. *Plant Physiol* **146**:515-528.

Datta SK (2005) Androgenic haploids: Factors controlling development and its application in crop improvement. *Curr Sci India* **89**:1870-1878.

Davies C, Wolf T and Robinson S (1999) Three putative sucrose transporters are differentially expressed in grapevine tissues. *Plant Sci* **147**:93-100.

Davik J and Honne B (2005) Genetic variance and breeding values for resistance to a wind-borne disease [*Sphaerotheca macularis* (Wallr. ex Fr.)] in strawberry (*Fragaria* × *ananassa* Duch.) estimated by exploring mixed and spatial models and pedigree information. *Theor Appl Genet* **111**:256-264.

Grillo G, Turi A, Licciulli F, Mignone F, Liuni S, Banfi S, Gennarino V, Horner D, Pavesi G, Picardi E and Pesole G (2010) UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Res* **38**:D75-D80.

Hackel A, Schauer N, Carrari F, Fernie AR, Grimm B and Kuhn C (2006) Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. *Plant J* **45**:180-192.

Henry C, Rabot A, Laloi M, Mortreau E, Sigogne M, Leduc N, Lemoine R, Sakr S, Vian A and Pelleschi-Travier S (2011) Regulation of RhSUC2, a sucrose transporter, is correlated with the light control of bud burst in *Rosa sp. Plant Cell Environ* **34**:1776-1789.

Hu S, Zhu X, Triggs JR, Tao Y, Wang Y, Lichtenstein L, Bissonnette M, Musch MW and Chang EB (2009) Inflammation-induced, 3'UTR-dependent translational inhibition of Hsp70 mRNA impairs intestinal homeostasis. *Am J Physiol Gastrointest Liver Physiol* **296**:G1003-1011.

Ito T, Motohashi R, Kuromori T, Noutoshi Y, Seki M, Kamiya A, Mizukado S, Sakurai T and Shinozaki K (2005) A resource of 5,814 *dissociation* transposon-tagged and sequence-indexed lines of *Arabidopsis* transposed from start loci on chromosome 5. *Plant Cell Physiol* **46**:1149-1153.

Jia H, Wang Y, Sun M, Li B, Han Y, Zhao Y, Li X, Ding N, Li C, Ji W and Jia W (2013a) Sucrose functions as a signal involved in the regulation of strawberry fruit development and ripening. *New Phytol* **198**:453-465.

Jia H-f, Lu D, Sun J-h, Li C-l, Xing Y, Qin L and Shen Y-y (2013b) Type 2C protein phosphatase ABI1 is a negative regulator of strawberry fruit ripening. *J Exp Bot* **64**:1677-1687.

Kim SW, Kim HJ, Kim JH, Kwon YK, Ahn MS, Jang YP and Liu JR (2011) A rapid, simple method for the genetic discrimination of intact *Arabidopsis thaliana* mutant seeds using metabolic profiling by direct analysis in real-time mass spectrometry. *Plant Methods* **7**:14.

Kuhn C, Hajirezaei MR, Fernie AR, Roessner-Tunali U, Czechowski T, Hirner B and Frommer WB (2003) The sucrose transporter *StSUT1* localizes to sieve elements in potato tuber phloem and influences tuber physiology and development. *Plant Physiol* **131**:102-113.

Li C, Jia H, Chai Y and Shen Y (2011) Abscisic acid perception and signaling transduction in strawberry: a model for non-climacteric fruit ripening. *Plant Signal Behav* **6**:1950-1953.

Martínez Zamora MG, Castagnaro AP and Díaz Ricci JC (2008) Genetic diversity of pto-like serine/threonine kinase disease resistance genes in cultivated and wild strawberries. *J Mol Evol* **67**:211-221.

Mayr C and Bartel D (2009) Widespread Shortening of 3′ UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell* **138**:673-684.

Mysore KS, Nam J and Gelvin SB (2000) An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. *Proc Natl Acad Sci U S A* **97**:948-953.

Reinders A, Sun Y, Karvonen KL and Ward JM (2012) Identification of amino acids important for substrate specificity in sucrose transporters using gene shuffling. *J Biol Chem* **287**:30296-30304.

Salvatierra A, Pimentel P, Moya-León M and Herrera R (2013) Increased accumulation of anthocyanins in *Fragaria chiloensis* fruits by transient suppression of *FcMYB1* gene. *Phytochemistry* **90**:25-36.

Sauer N (2007) Molecular physiology of higher plant sucrose transporters. *FEBS Lett* **581**:2309-2317.

Schneidereit A, Imlau A and Sauer N (2008) Conserved *cis*-regulatory elements for DNA-binding-with-one-finger and homeo-domain-leucine-zipper

transcription factors regulate companion cell-specific expression of the *Arabidopsis thaliana SUCROSE TRANSPORTER 2* gene. *Planta* **228**:651-662.

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arus P, Dandekar AM, Lewers K, Brown SK, Davis TM, Gardiner SE, Potter D and Veilleux RE (2008) Multiple models for Rosaceae genomics. *Plant Physiol* **147**:985-1003.

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton JM, Rees DJ, Williams KP, Holt SH, Ruiz Rojas JJ, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman TL, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Jr., Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Lopez Girona E, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE and Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). *Nat Genet* **43**:109-116.

Srivastava AC, Ganesan S, Ismail IO and Ayre BG (2008) Functional characterization of the Arabidopsis AtSUC2 Sucrose/H+ symporter by tissue-specific complementation reveals an essential role in phloem loading but not in long-distance transport. *Plant Physiol* **148**:200-211.

Tanguay RL and Gallie DR (1996) The effect of the length of the 3'-untranslated region on expression in plants. *FEBS Lett* **394**:285-288.

Thill J, Miosic S, Gotame T, Mikulic-Petkovsek M, Gosch C, Veberic R, Preuss A, Schwab W, Stampar F, Stich K and Halbwirth H (2013) Differential expression of flavonoid 3′ -hydroxylase during fruit development establishes the different B-ring hydroxylation patterns of flavonoids in *Fragaria × ananassa* and *Fragaria vesca*. *Plant Physiol Bioch*.

Veilleux RE, Mills KP, Baxter AJ, Upham KT, Ferguson TJ, Holt SH, Lu N, Ruiz-Rojas JJ, Pantazis CJ, Davis CM, Lindsay RC, Powell FL, Dan Y, Dickerman AW, Oosumi T and Shulaev V (2012) Transposon tagging in diploid strawberry. *Plant Biotechnol J* **10**:985-994.

Vellicce G, Ricci J, Hernández L and Castagnaro AP (2006) Enhanced Resistance to *Botrytis cinerea* Mediated by the Transgenic Expression of the Chitinase Gene *ch5B* in Strawberry. *Transgenic Res* **15**:57-68.

Weise A, Lalonde S, Kühn C, Frommer WB and Ward JM (2008) Introns control expression of sucrose transporter LeSUT1 in trichomes, companion cells and in guard cells. *Plant Mol Biol* **68**:251-262.

Wilkie GS, Dickson KS and Gray NK (2003) Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends Biochem Sci* **28**:182-188.

Williams LE, Lemoine R and Sauer N (2000) Sugar transporters in higher plants--a diversity of roles and complex regulation. *Trends Plant Sci* **5**:283-290.

Wippel K and Sauer N (2012) *Arabidopsis* SUC1 loads the phloem in *suc2* mutants when expressed from the *SUC2* promoter. *J Exp Bot* **63**:669-679.

Xiong Y, DeFraia C, Williams D, Zhang X and Mou Z (2009) Characterization of Arabidopsis 6-phosphogluconolactonase T-DNA insertion mutants reveals an essential role for the oxidative section of the plastidic pentose phosphate pathway in plant growth and development. *Plant Cell Physiol* **50**:1277-1291.

Yu D, Tang H, Zhang Y, Du Z, Yu H and Chen Q (2012) Comparison and improvement of different methods of RNA isolation from strawberry (*Fragria* x *ananassa*). *J Agric Sci* **4**:51-56.

Zainal N and Turnbull C (2011) Short Communication: Effects of different light conditions on various *CONSTANS-LIKE* mutants and their response to flowering time. *AsPac J Mol Biol Biotechnol* **19**.

Zhang J, Yu C, Krishnaswamy L and Peterson T (2011) Transposable elements as catalysts for chromosome rearrangements. *Methods Mol Biol* **701**:315-326.

Zhao Y, Liu Q and Davis RE (2004) Transgene expression in strawberries driven by a heterologous phloem-specific promoter. *Plant Cell Rep* **23**:224-230.

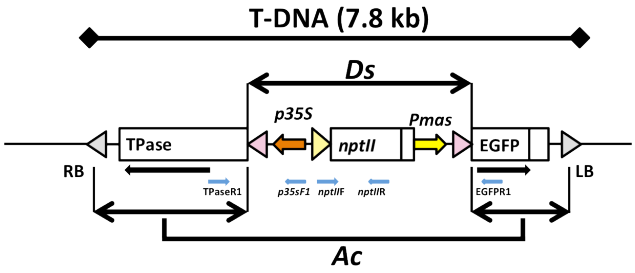


Figure 19 Schematic map of construct used for transformation of the diploid strawberry population that gave rise to phenotypic mutant LP119. RB, right border; TPase, maize transposase; *Ds*, *Dissociator* part of the element; P35S, cauliflower mosaic virus promoter; *NPTII*, neomycin phosphotransferase gene; Pmas, mannopine synthase promoter; EGFP, enhanced green fluorescent protein gene; LB, left border; triangular borders of the *Ds* element represent inverted repeats. The *Activator* (*Ac*) part of the construct occurs on both sides of the *Ds* element. The T-DNA was inserted into plasmid psKI074. Primer positions were indicated by blue arrows: FDS (TPaseR1 and p35sF1), *NPTII* (NPTIIF and NPTIIR), EDS (TPaseR1 and EGFPR1).

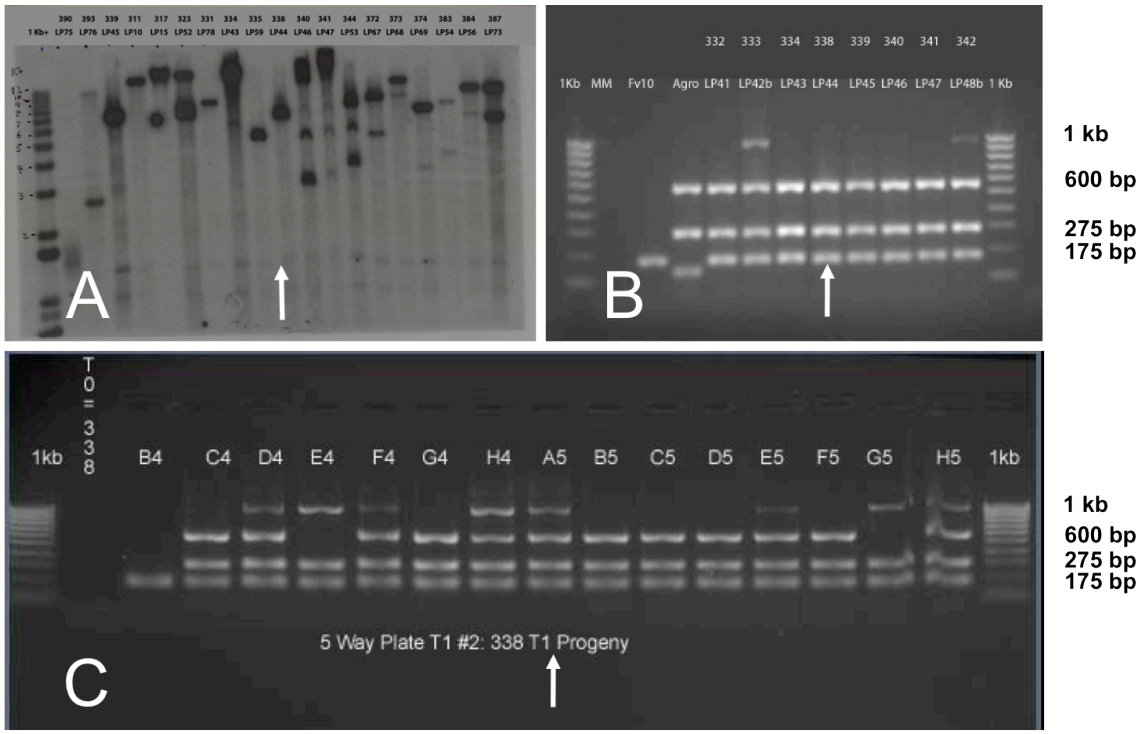


Figure 20 Southern blot and multiplex PCR in T_0 and T_1 plants. A) Southern blot using FDS probe showing single T-DNA insertion for LP119 (also called LP44, white arrow); B) Multiplex PCR of T_0 plants after transformation with the *pAcDs* construct showing EDS (1 kb), FDS (600 bp), *NPTII* (275 bp) and *FvAAT3* (175 bp). White arrow showed T_0 plant LP119; C) Multiplex PCR of T_1 plants using same primer sets. White arrow showed T_1 mutant (A5) with increased anthocyanin concentration that was derived from LP119 after self-pollination.

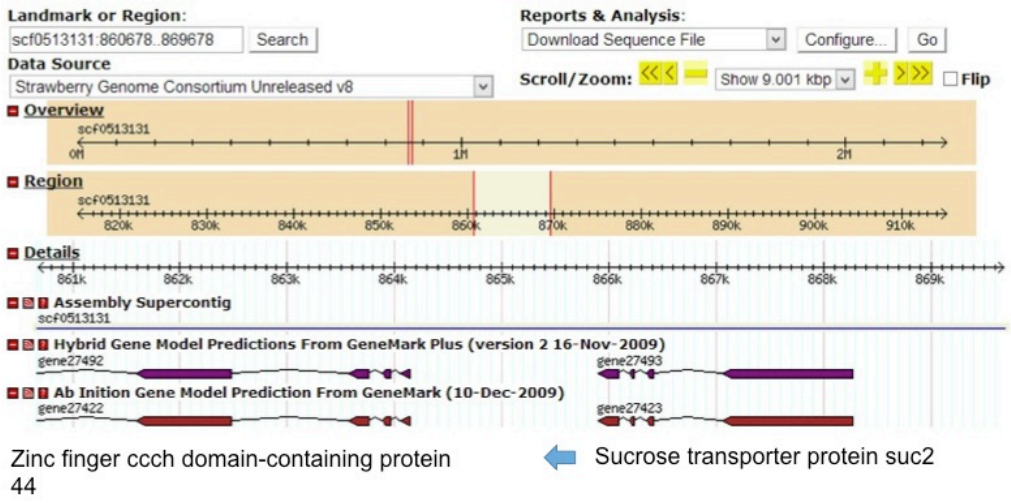


Figure 21 T-DNA insertion site of T_0-LP119 as shown in genome browser (blue arrow – direction showed the beginning and end of the HiTAIL-PCR sequence derived from Tpase-TAIL). In hybrid gene model, the insertion was 226 bp from stop codon of gene27493, sucrose transporter protein-2, *FvSUC2*; and 1.5 kb from starting codon of gene27492, zinc finger CCCH domain-containing protein 44.

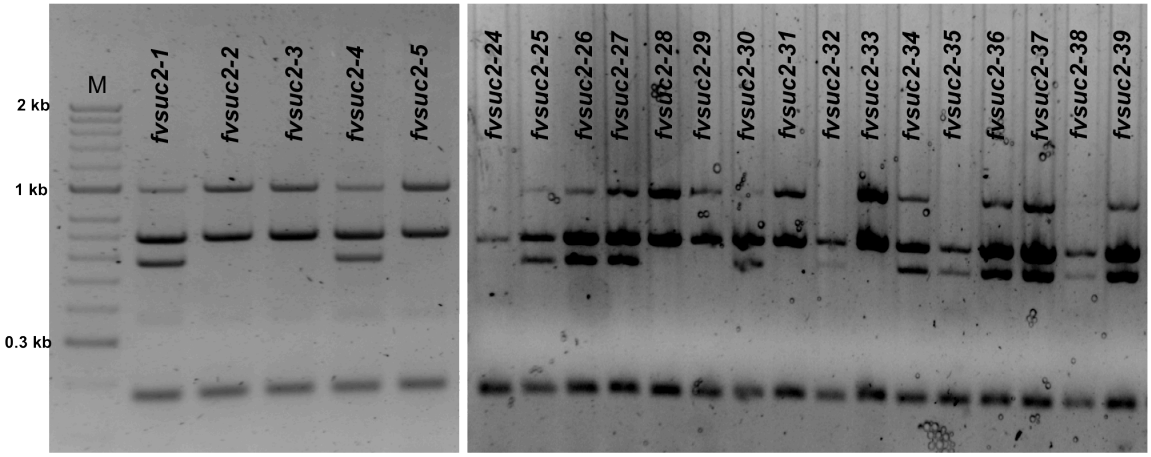


Figure 22 Multiplex PCR of 21 T_2 plants derived from self-pollination of T_1 mutant LP119-A5. The 1 kb product was EDS (TPaseR1/EGFPR1), the 800 bp product was *NPTII* (NPTIIF2/NPTIIR2), the 600 bp product was FDS (TPaseR1/p35sF1), and the 175 bp band was *FvAAT3*. All mutants had *NPTII* while EDS and FDS segregated among them. None of them were wild type, suggesting homozygote of T-DNA insertion. Gel image was inverted to white background for better quality. M, HypperLadder™ II.

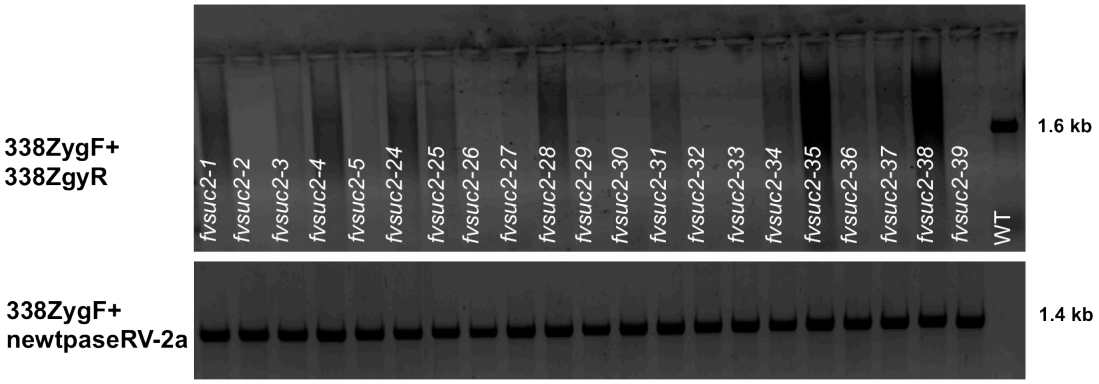


Figure 23 Zygosity test of T_2 plants. A) PCR with 338ZygF and 338ZygR, primers designed from sequences flanking T-DNA insertion site in strawberry genome. Amplification was found only in wild type plants. (lane 22) B) PCR with 338ZygF and newtpaseRV-2a. Amplifications were found in all 21 mutant plants (lanes 1-21) but not in wild type control (lane 22). Gel image was inverted for better quality.

Figure 24 Phenotype of T_2 plants homozygous for *AcDs* T-DNA insertion 226 bp from the 3' end of the sucrose transporter gene 21 weeks after germination, after growing in a growth chamber for 18 weeks then transferred to the greenhouse for 3 weeks. A) T_2 mutant in a 10 cm pot. B) Mutant (right) and wild type plant (left), both growing in 10 cm pots. C) Trifoliate leaf from wild type plant (left) and mutant (right); note the increased anthocyanin pigmentation in the primary vein and petiolule of mutant leaf. D) Fruit at different developmental stages of mutant (right) and wild type (left). Fruit at same row were at the same development stage.

Table 8 Multiplex PCR results of characterizing 21 T_2 plants. Seven plants had EDS and *NPTII*, eight had EDS, FDS and *NPTII* while six plants had FDS and *NPTII*.

PCR results	T_2 plant ID	Number of plants
EDS, *NPTII*	*fvsuc2-2,3,5,28,29,31,33*	7
EDS, FDS, *NPTII*	*fvsuc2-1,4,26,27, 34,36,37,39*	8
FDS, *NPTII*	*fvsuc2-24,25,30,32,35,38*	6

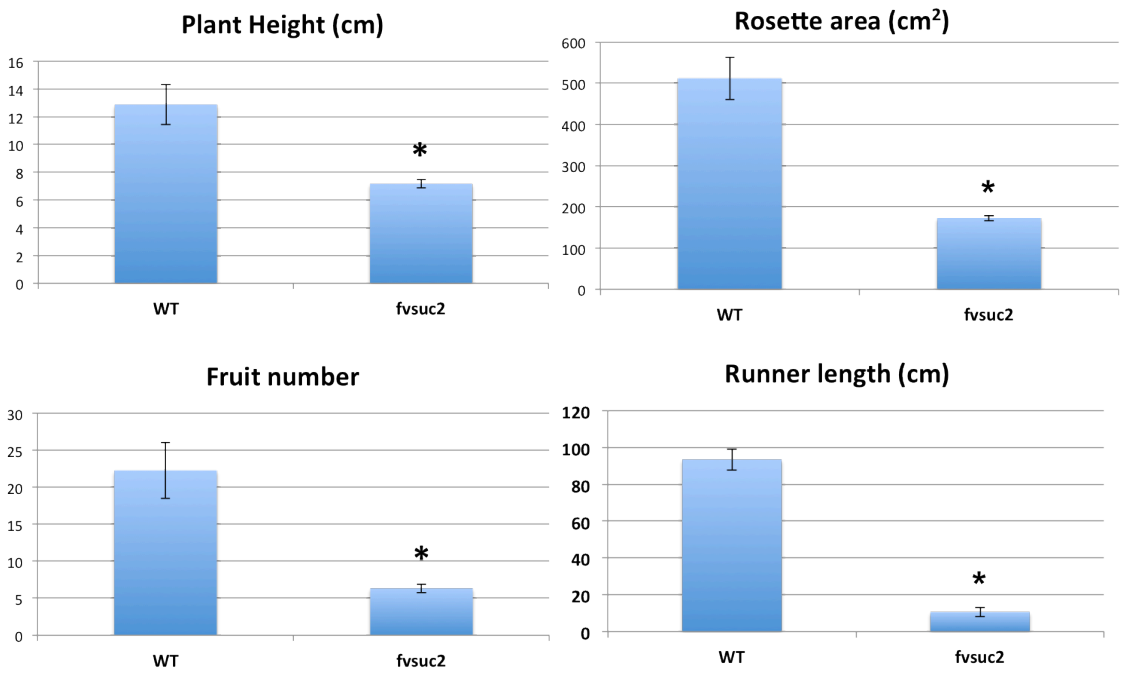


Figure 25 Phenotypic data of *AcDs* mutants and wild type plants. Data were collected 6 weeks after moving plants from growth chamber to greenhouse conditions. Error bar represented standard error of the mean. Data represented mean values of at least three replicates. p-value was calculated using Student's t-test. * indicates significant p-values ($p<0.01$).

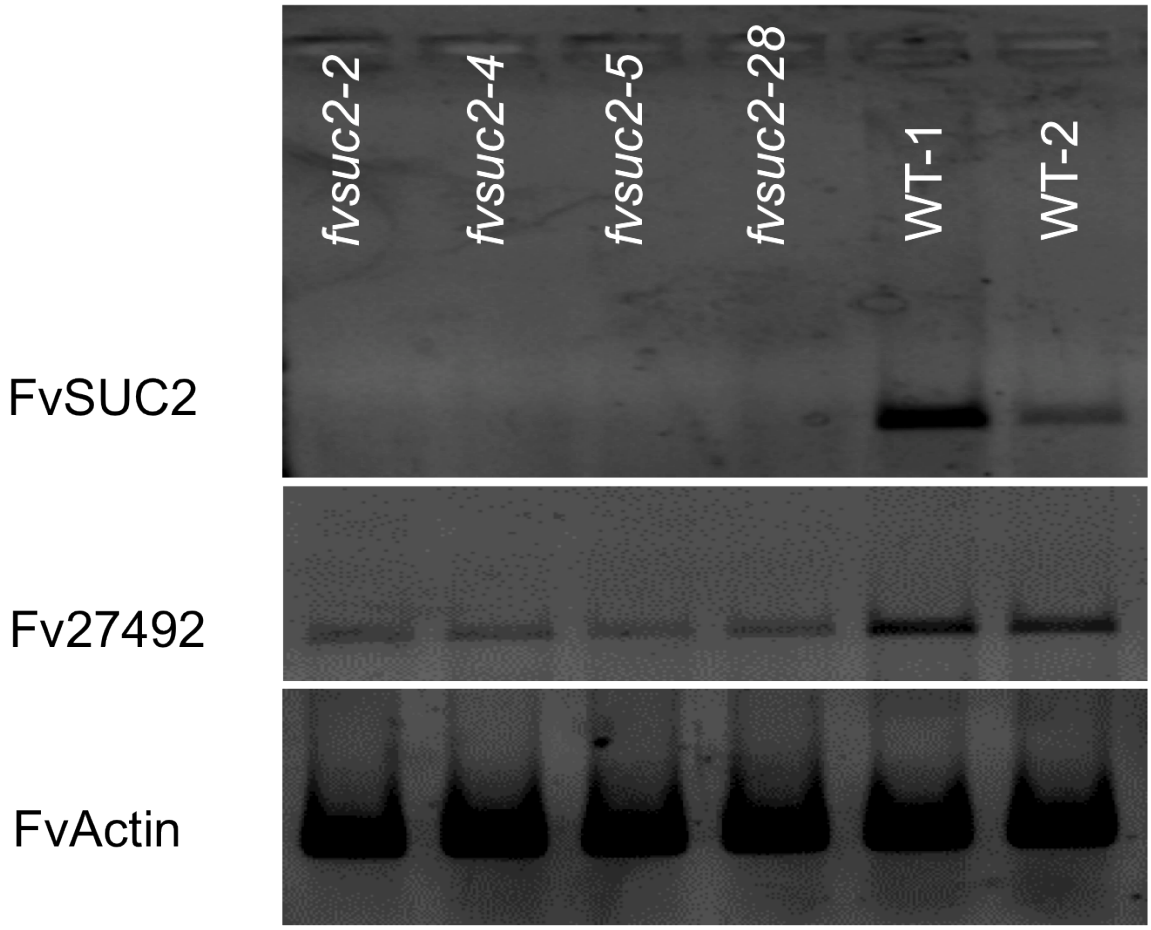


Figure 26 Semi-quantitative RT-PCR of genes *FvSUC2* and *Fv27492* in leaf tissue of four mutant (*fvsuc2-2*, *fvsuc2-4*, *fvsuc2-5*, *fvsuc2-28*) and two wild type plants (WT-1, WT-2). *FvActin* was used as control. Gel image was inverted for better quality.

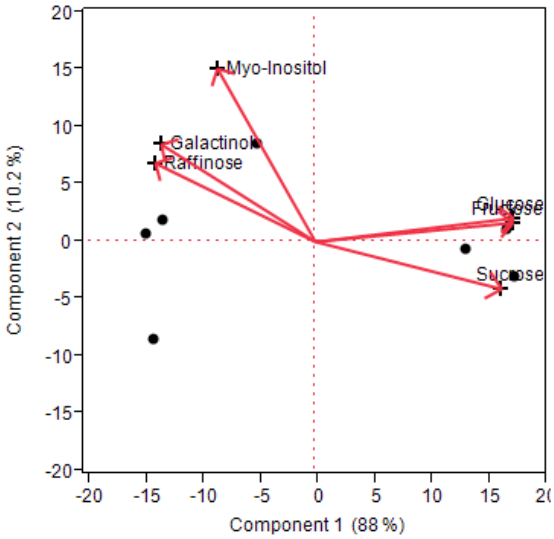


Figure 27 Biplots of principal component analysis of six sugar compounds (Myo-inositol, galactinol2, raffinose, glucose, fructose and sucrose) in wild type (indicated by four dots on left) and mutants (indicated by dots on right). Each red arrow represented one compound and each dot represented one sample.

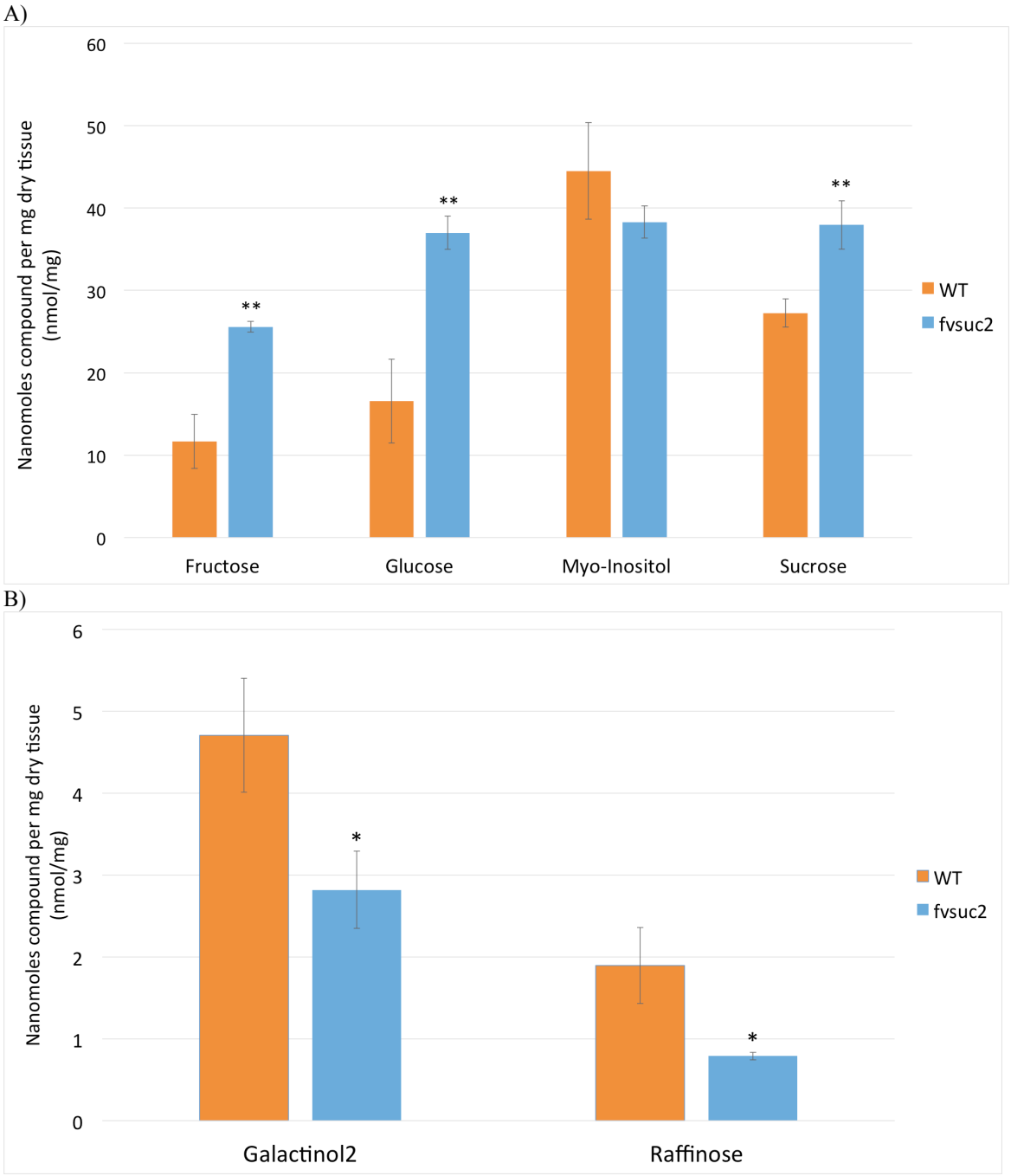


Figure 28 GCMS analysis of carbohydrate concentrations in leaf extracts of *fvsuc2* and wild type strawberry plants. A) Fructose, glucose, myo-inositol and sucrose contents; B) trisaccharides galactinol2 and raffinose contents. p-values were calculated using Student's t-test. (*) indicates significant p values ($p<0.01$), (**) indicates highly significant p values ($p<0.002$).

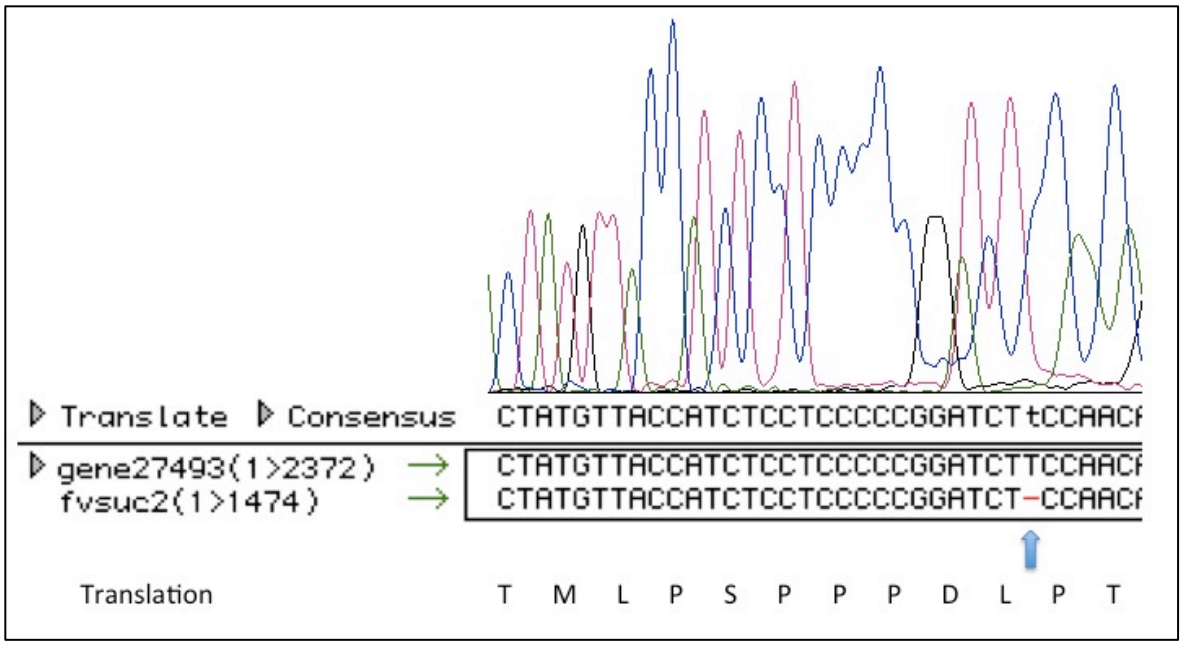


Figure 29 Sequence alignment of the *FvSUC2* gene from strawberry genome browser with that obtained from the *fvsuc2* mutant revealed a one-nucleotide deletion in exon 4. Blue arrow indicates deleted nucleotide.

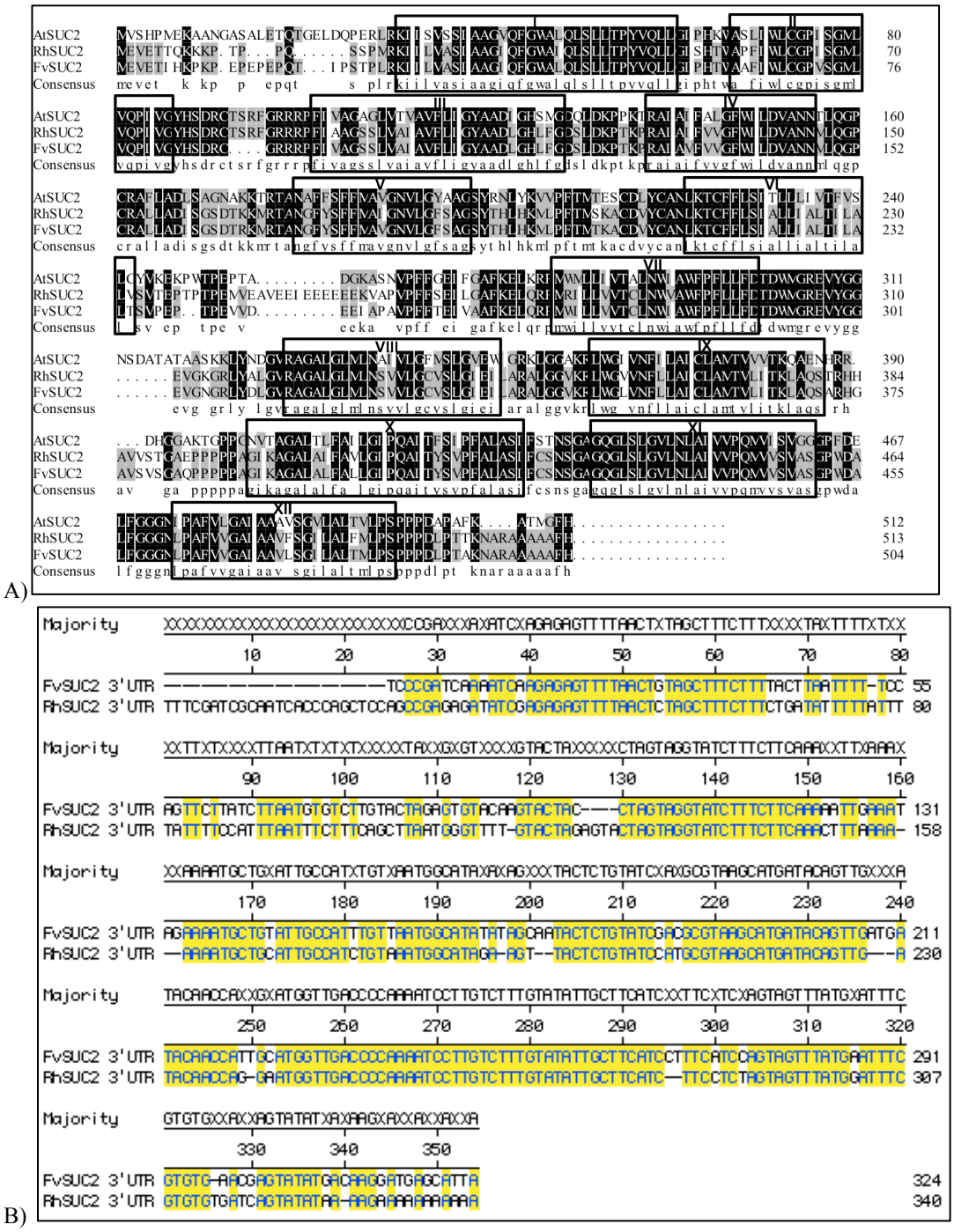


Figure 30 mRNA sequences and 3'UTR sequences alignment of *FvSUC2* with other orthologs. A. Amino acid sequence alignment of *FvSUC2* with *AtSUC2* and *RhSUC2* using Clustal-W method. Black highlight indicated identical amino acids; gray shading indicated similar amino acid residues. I - XII were 12 proposed membrane-spanning domains according to (Davies et al. 1999; Srivastava et al. 2008). B. Nucleotide sequence alignment of partial *FvSUC2* 3'UTR sequences with *RhSUC2* 3'UTR sequences using CLUSTAL-W method with the Megalign tool of Lasergene 10 software suite. Highlighted part indicates similar sequence.

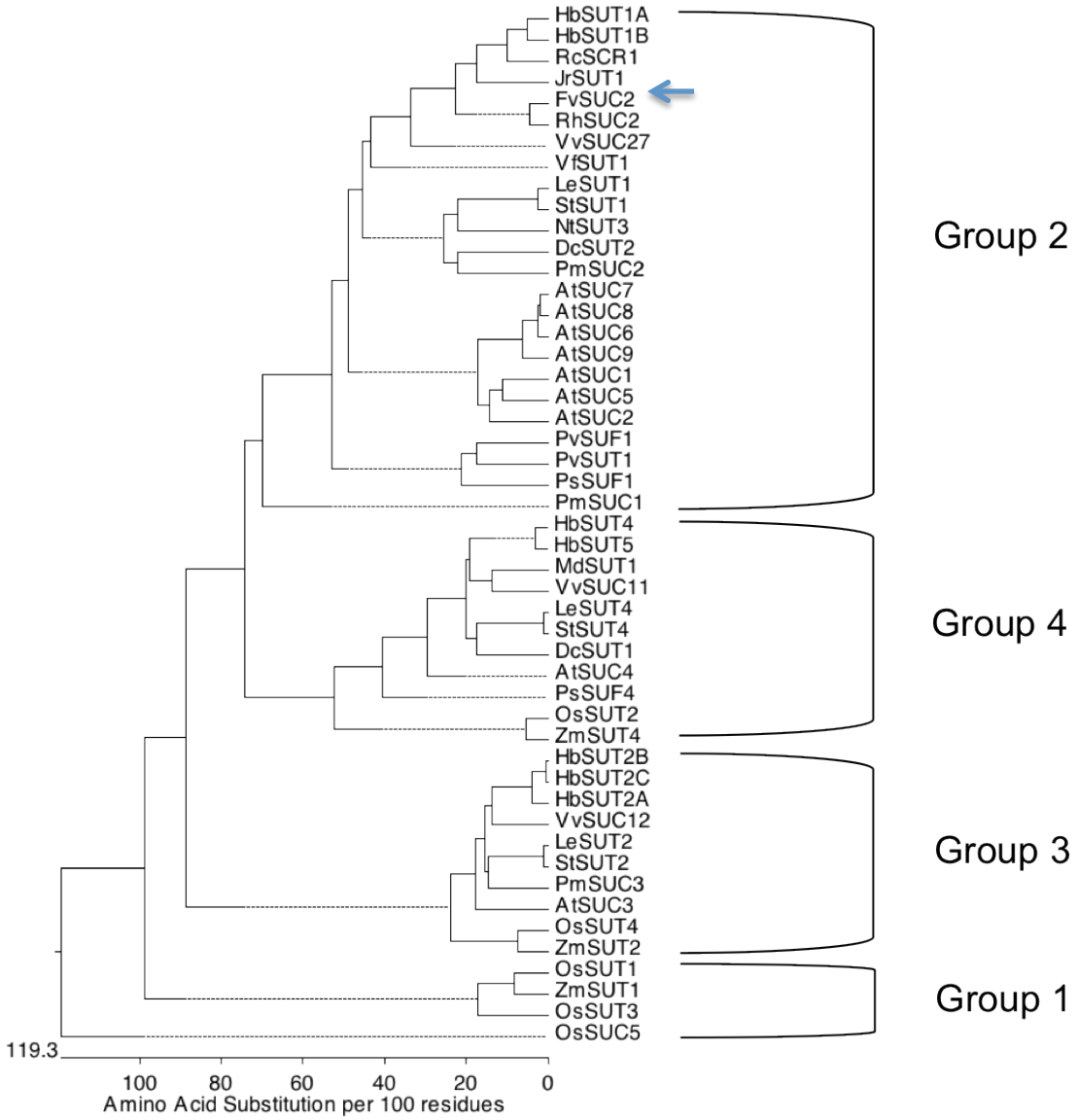
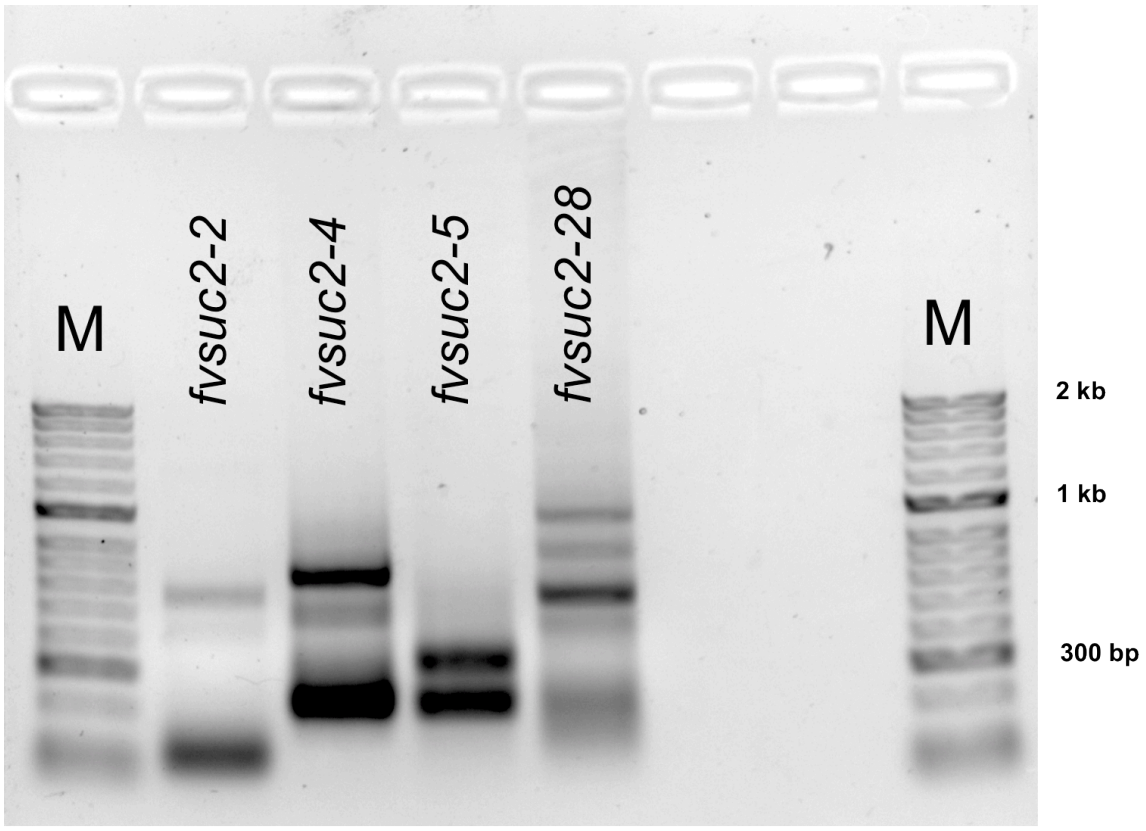


Figure 31 Phylogenetic tree of *FvSUC2* with other characterized plant sucrose transporter genes using CLUSAL-W method with Megalign tool in Lasergene 10 software suite. *FvSUC2* (indicated by arrow) was in the same group as *AtSUC2* and *RhSUC2*.

Supplemental material.

Supplemental Table 5 Primers used in this study and their purpose.

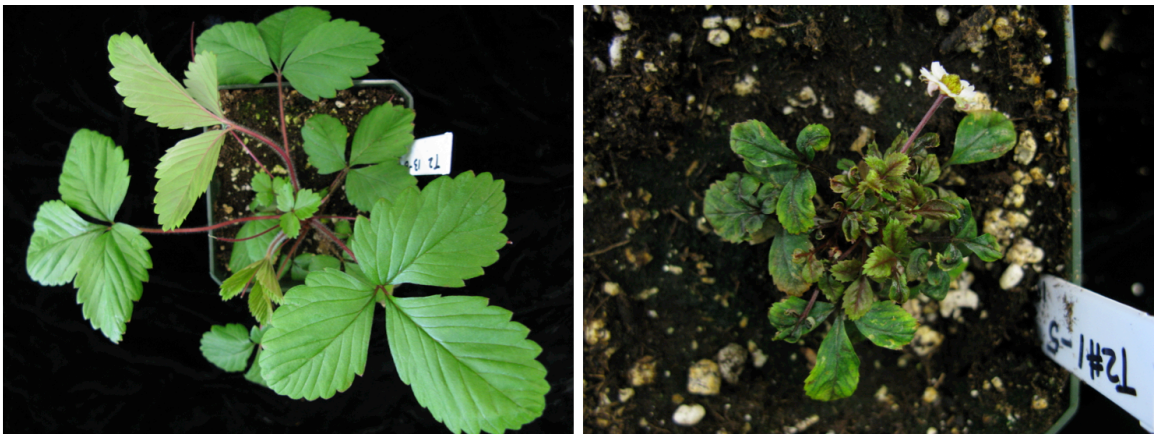
Primer name	Primer sequence (5'-3')	Purpose
TPaseR1	GGTGAAATGCTGCCATAC	Multiplex PCR
EGFPR1	TTGTACAGCTCGTCCATG	Multiplex PCR
p35sF1	ACGCACAATCCCACTATC	Multiplex PCR
NPTIIF2	CGGCTATGACTGGGCACAACAG	Multiplex PCR
NPTIIR2	GGCGATACCGTAAAGCACGAGG	Multiplex PCR
NPTIIF	TATGACTGGGCACAACAG	Multiplex PCR
NPTIIR	GCATCAGCCATGATGGATAC	Multiplex PCR
FvAAT3F	GTGACTTGGTTAACTTGCTC	Multiplex PCR
FvAAT3R	AAATTAGTCCAGCTCGTGAA	Multiplex PCR
Fv27493F	CCGGAAGTCGTGGATGAGGAGA	RT-PCR and sequencing
Fv27493R	TGCTAGAGCAAAGGGAACACTGTAGGT	RT-PCR and sequencing
Fv27492F	AAAGAAAGAAGAGATTTGCGAAGAGGTT	RT-PCR
Fv27492R	TTAATTTGCAAGAGCTGGTGAGAGTTC	RT-PCR
newtpaseRV-2a	GTTAGCCAAGAGCCCAAGACTTATCAC	Zygosity test
FvSUC2F	ATGGAAGTCGAAACCATCCACAAG	Sequencing
FvSUC2R	TCAATGGAATGCAGCTGCTGC	Sequencing
338zygF	TCAATCTCGAAACCCATCCTCCAT	Zygosity test
338zygR	ACACATGAAGAACAATGCTGTAACTGACC	Zygosity test
FvActinF	GGTCTCGAACATTATCTGGGTCAT	RT-PCR
FvActinR	AGGCCGGGTTTGCTGGAG	RT-PCR



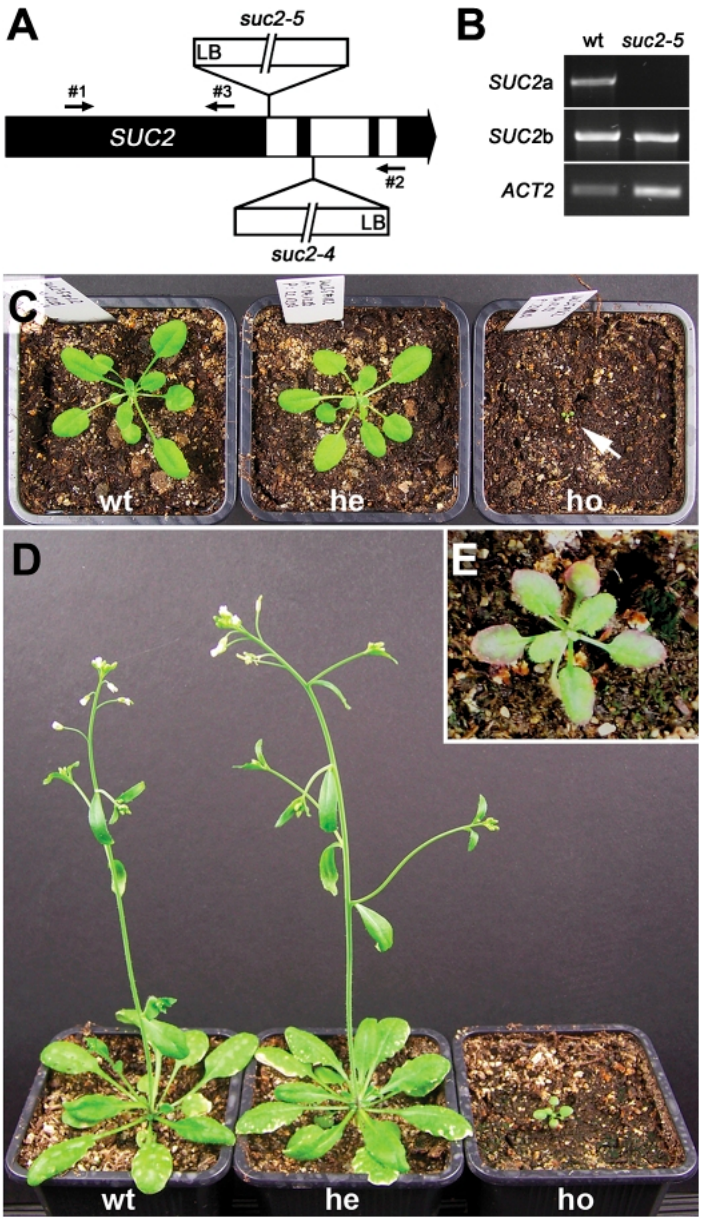
Supplemental Figure 3 *Ds*5 HiTAIL-PCR results of four T_2 mutants. *Ds*5 primers set were used to amplify flanking sequences of reinserted *Ds* from *fvsuc2-2/4/5/28* (lanes 2-5). Selected products were purified and sequenced before performing BLAST search in genome browser to locate insertion sites. M, HyperLadder™ II. Gel image was inverted to white background for better quality.

Supplemental Table 6 Insertion sites of reinserted *Ds* from four T_2 mutants in strawberry genome compared with T_0 launch pad location, as well as the genomic region of reinsertion sites.

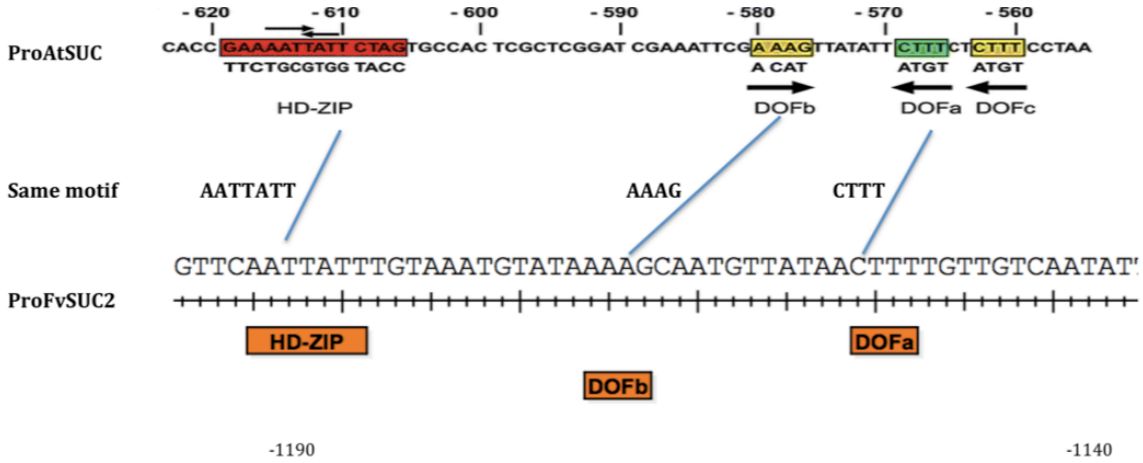
ID	**Original insertion site**	***Ds* reinsertion site**	**Genomic loci of *Ds***	**Candidate gene**	**Gene function**
LP119	chr02:13951314		3'UTR of gene27493	*FvSUC2*	Sugar transport
fvsuc2-2		chr02:31941485	Intron of gene16106	UPF0558 protein C1orf156 homolog	Unknown
fvsuc2-4		chr06:36637124	Intergenic region	NA	NA
fvsuc2-5		chr02:31941485	Intron of gene16106	UPF0558 protein C1orf156 homolog	Unknown
fvsuc2-28		chr02:31941485	Intron of gene16106	UPF0558 protein C1orf156 homolog	Unknown



Supplemental Figure 4 Picture of *fvsuc2-5* (right) and wild type (left) plants 16 weeks after germination in growth chamber conditions. Mutant shows accumulated anthocyanin and dwarf phenotype. Plants were grown in 10 cm pots.



Supplemental Figure 5 Arabidopsis knockout mutant phenotype of the *AtSUC2* gene. The T-DNA homozygote mutant showed stunted phenotype with increased anthocyanin content while hemizygous mutant did not show obvious phenotype compared to wild type plants (C, D, E). Picture is from (Wippel and Sauer 2012).



Supplemental Figure 6 Alignment of promoter sequences from *AtSUC2* (top, picture from (Schneidereit et al. 2008)) and *FvSUC2* (bottom, predicted from genome browser). Similar motifs representing DNA-binding domains were shown in the middle. Numbers on the top and bottom of sequences indicates relative position from beginning of transcript.